


11/08/00
JC715 U.S. PTO

09710633 11/08/00

11.13 00

Please type a plus sign (+) inside this box  Approved for use through 10/31/2002. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

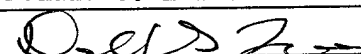
JC025 U.S. PTO
09/710633
11/08/00

UTILITY PATENT APPLICATION TRANSMITTAL (Only for new nonprovisional applications under 37 CFR 1.53(b))	Attorney Docket No. TSRI 478.0C1	
	First Inventor Kent, et al.	
	Title	SYNTHESIS OF PROTEINS BY NATIVE.
	Express Mail Label No. EL587956875US	

APPLICATION ELEMENTS See MPEP chapter 600 concerning utility patent application contents.	ADDRESS TO: Assistant Commissioner for Patents Box Patent Application Washington, DC 20231
1. <input checked="" type="checkbox"/> Fee Transmittal Form (e.g., PTO/SB/17) (Submit an original and a duplicate for fee processing)	7. <input type="checkbox"/> CD-ROM or CD-R in duplicate, large table or Computer Program (Appendix)
2. <input checked="" type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27.	8. Nucleotide and/or Amino Acid Sequence Submission (if applicable, all necessary)
3. <input checked="" type="checkbox"/> Specification [Total Pages 59] (preferred arrangement set forth below) <ul style="list-style-type: none">- Descriptive title of the invention- Cross Reference to Related Applications- Statement Regarding Fed sponsored R & D- Reference to sequence listing, a table, or a computer program listing appendix- Background of the Invention- Brief Summary of the Invention- Brief Description of the Drawings (if filed)- Detailed Description- Claim(s)- Abstract of the Disclosure	a. <input type="checkbox"/> Computer Readable Form (CRF) b. Specification Sequence Listing on: <ul style="list-style-type: none">i. <input type="checkbox"/> CD-ROM or CD-R (2 copies); orii. <input type="checkbox"/> paper c. <input type="checkbox"/> Statements verifying identity of above copies
4. <input type="checkbox"/> Drawing(s) (35 U.S.C. 113) [Total Sheets]	ACCOMPANYING APPLICATION PARTS 9. <input type="checkbox"/> Assignment Papers (cover sheet & document(s)) 10. <input type="checkbox"/> 37 CFR 3.73(b) Statement (when there is an assignee) <input type="checkbox"/> Power of Attorney 11. <input type="checkbox"/> English Translation Document (if applicable) 12. <input type="checkbox"/> Information Disclosure Statement (IDS)/PTO-1449 <input type="checkbox"/> Copies of IDS Citations 13. <input checked="" type="checkbox"/> Preliminary Amendment 14. <input checked="" type="checkbox"/> Return Receipt Postcard (MPEP 503) (Should be specifically itemized) 15. <input type="checkbox"/> Certified Copy of Priority Document(s) (if foreign priority is claimed) 16. <input checked="" type="checkbox"/> Other: Certificate of Express Mail
5. Oath or Declaration [Total Pages 4] a. <input type="checkbox"/> Newly executed (original or copy) b. <input checked="" type="checkbox"/> Copy from a prior application (37 CFR 1.63 (d)) (for continuation/divisional with Box 17 completed) i. <input type="checkbox"/> DELETION OF INVENTOR(S) Signed statement attached deleting inventor(s) named in the prior application, see 37 CFR 1.63(d)(2) and 1.33(b).	
6. <input type="checkbox"/> Application Data Sheet. See 37 CFR 1.76	
17. If a CONTINUING APPLICATION, check appropriate box, and supply the requisite information below and in a preliminary amendment, or in an Application Data Sheet under 37 CFR 1.76: <input checked="" type="checkbox"/> Continuation <input type="checkbox"/> Divisional <input type="checkbox"/> Continuation-in-part (CIP) of prior application No. 08, 945,997 Prior application information Examiner J. Russel Group / Art Unit 1653	

For CONTINUATION OR DIVISIONAL APPS only: The entire disclosure of the prior application, from which an oath or declaration is supplied under Box 5b, is considered a part of the disclosure of the accompanying continuation or divisional application and is hereby incorporated by reference. The incorporation can only be relied upon when a portion has been inadvertently omitted from the submitted application parts.

18. CORRESPONDENCE ADDRESS					
<input type="checkbox"/> Customer Number or Bar Code Label			or <input checked="" type="checkbox"/> Correspondence address below		
(Insert Customer No. or Attach bar code label here)					
Name	THE SCRIPPS RESEARCH INSTITUTE				
Address	10550 North Torrey Pines Road				
	Mail Drop: TPC-8				
City	La Jolla	State	CA	Zip Code	92037
Country	US	Telephone	(858) 784-2937	Fax	858-784-9399

Name (Print/Type)	Donald G. Lewis	Registration No. (Attorney/Agent)	28,636
Signature		Date	11/8/00

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Box Patent Application, Washington, DC 20231

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

FEE TRANSMITTAL for FY 2001

Patent fees are subject to annual revision.

Complete if Known

Application Number	
Filing Date	November 8, 2000
First Named Inventor	Kent
Examiner Name	
Group Art Unit	
Attorney Docket No.	TSRI 478.0C1

TOTAL AMOUNT OF PAYMENT (\$ 526.00)

METHOD OF PAYMENT

- 1.
- ☒
- The Commissioner is hereby authorized to charge indicated fees and credit any overpayments to:

Deposit Account Number	19-0962
Deposit Account Name	

☒ Charge Any Additional Fee Required Under 37 CFR 1.16 and 1.17☒ Applicant claims small entity status. See 37 CFR 1.27

- 2.
- ☒
- Payment Enclosed:

☒ Check ☐ Credit card ☐ Money Order ☐ Other**FEE CALCULATION****1. BASIC FILING FEE**

Large Entity Fee Code (\$)	Small Entity Fee Code (\$)	Fee Description	Fee Paid
101 710	201 355	Utility filing fee	355
106 320	206 160	Design filing fee	
107 490	207 245	Plant filing fee	
108 710	208 355	Reissue filing fee	
114 150	214 75	Provisional filing fee	

SUBTOTAL (1) (\$ 355.00)

2. EXTRA CLAIM FEES

Total Claims	Extra Claims	Fee from below	Fee Paid
24	-20** = 4	9	36
3	-3** = 0		
Multiple Dependent		135	135

Large Entity Fee Code (\$)	Small Entity Fee Code (\$)	Fee Description
103 18	203 9	Claims in excess of 20
102 80	202 40	Independent claims in excess of 3
104 270	204 135	Multiple dependent claim, if not paid
109 80	209 40	** Reissue independent claims over original patent
110 18	210 9	** Reissue claims in excess of 20 and over original patent

SUBTOTAL (2) (\$ 171.00)

**or number previously paid, if greater; For Reissues, see above

FEE CALCULATION (continued)**3. ADDITIONAL FEES**

Large Entity Fee Code (\$)	Small Entity Fee Code (\$)	Fee Description	Fee Paid
105 130	205 65	Surcharge - late filing fee or oath	
127 50	227 25	Surcharge - late provisional filing fee or cover sheet	
139 130	139 130	Non-English specification	
147 2,520	147 2,520	For filing a request for <i>ex parte</i> reexamination	
112 920*	112 920*	Requesting publication of SIR prior to Examiner action	
113 1,840*	113 1,840*	Requesting publication of SIR after Examiner action	
115 110	215 55	Extension for reply within first month	
116 390	216 195	Extension for reply within second month	
117 890	217 445	Extension for reply within third month	
118 1,390	218 695	Extension for reply within fourth month	
128 1,890	228 945	Extension for reply within fifth month	
119 310	219 155	Notice of Appeal	
120 310	220 155	Filing a brief in support of an appeal	
121 270	221 135	Request for oral hearing	
138 1,510	138 1,510	Petition to institute a public use proceeding	
140 110	240 55	Petition to revive - unavoidable	
141 1,240	241 620	Petition to revive - unintentional	
142 1,240	242 620	Utility issue fee (or reissue)	
143 440	243 220	Design issue fee	
144 600	244 300	Plant issue fee	
122 130	122 130	Petitions to the Commissioner	
123 50	123 50	Petitions related to provisional applications	
126 240	126 240	Submission of Information Disclosure Stmt	
581 40	581 40	Recording each patent assignment per property (times number of properties)	
146 710	246 355	Filing a submission after final rejection (37 CFR § 1.129(a))	
149 710	249 355	For each additional invention to be examined (37 CFR § 1.129(b))	
179 710	279 355	Request for Continued Examination (RCE)	
169 900	169 900	Request for expedited examination of a design application	

Other fee (specify) _____

* Reduced by Basic Filing Fee Paid

SUBTOTAL (3) (\$)

SUBMITTED BY**Complete (if applicable)**

Name (Print/Type)	Donald G. Lewis	Registration No. (Attorney/Agent)	28,636	Telephone	(858) 784-2937
Signature	<i>Donald G. Lewis</i>	Date	11/8/00		

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Washington, DC 20231 DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

11/08/00
JC715 U.S. PTO

11-13-00

A

PATENT
IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATION UNDER 37 CFR 1.10

JC715 U.S. PTO
09/710633
11/08/00

I hereby certify that this UTILITY PATENT APPLICATION TRANSMITTAL and the documents referred to therein are being deposited with the United States Postal Service "Express Mail Post Office to Address" Mail Label No. EL587945875US under 37 CFR 1.10 on the date indicated below and is addressed to: Box Patent Application, Assistant Commissioner for Patents, Washington, D.C. 20231.

Donald G. Lewis
Donald G. Lewis, Reg. No. 28,636

Nov 8, 2000
Date of Deposit

Applicant: Kent, et al.)	
)	
Serial No.: Unassigned)	Group Art Unit: Unassigned
)	
Filed: November 8, 2000)	Examiner: Unassigned
)	
Title: SYNTHESIS OF PROTEINS BY NATIVE CHEMICAL LIGATION)	Our Ref.: TSRI 478.0C1
)	

09710633 110800

Figure 1 consists of 12 subplots arranged in a 4x3 grid, labeled (a) through (l). Each subplot shows the probability distribution $P(x)$ on the y-axis (ranging from 0 to 0.02) versus the position x on the x-axis (ranging from 0 to 100). The subplots are labeled (a) through (l). The distributions show a peak that moves and changes shape over time. The peak is located at $x=0$ in (a), (c), (e), (g), (i), and (k), and at $x=100$ in (b), (d), (f), (h), (j), and (l). The peak height increases from 0.01 in (a) to 0.02 in (l). The distributions are symmetric in (a), (c), (e), (g), (i), and (k), and asymmetric in (b), (d), (f), (h), (j), and (l).

CERTIFICATION UNDER 37 CFR 1.10


Donald G. Lewis, Reg. No. 28,636

Nov 8, 2000
Date of Deposit

PRELIMINARY AMENDMENT

ATTN: BOX PATENT APPLICATION

Prior to examination of the above-identified patent application on the merits, which is a Continuation application of pending prior application U.S. Serial No. 08/945,997 filed under 37 CFR 1.53(b) concurrently herewith, please enter the following amendments.

Please add new claims 8-31.

--8. A method for producing a desired protein or domain thereof, which comprises admixing:

(I) a first oligopeptide, said first oligopeptide comprising a fragment of said desired protein or domain thereof, and having a C-terminal thioester; and

(II) a second oligopeptide, said second oligopeptide comprising a fragment of said desired protein or domain thereof, and having an N-terminal amino acid residue having an unoxidized sulfhydryl side chain and a free amino group that is capable of forming a β -aminothioester linkage with said C-terminal thioester that rearranges to form an amide bond therein between; wherein said admixing is conducted under conditions sufficient to permit the formation of an amide bond between the C-terminus of said first oligopeptide fragment and the N-terminus of said second oligopeptide fragment.

9. The method of claim 8, wherein said N-terminal amino acid residue is a cysteine residue.

10. The method of claim 8, wherein said desired protein is a naturally isolatable protein.

11. The method of claim 8, wherein said desired protein is a derivative of a naturally isolatable protein that contains one or more cysteine residues that are not found in said naturally isolatable protein.

12. The method of any of claims 10 or 11, wherein said protein is a mammalian protein.

13. The method of claim 12, wherein said mammalian protein is a human protein.

14. The method of claim 13, wherein said human protein is a cytokine.

09710631-100000

15. The method of claim 8, wherein in said method said second oligopeptide has a C-terminal thioacid, and wherein said method additionally comprises the steps of:

(A) converting said thioacid to a thioester; and

(B) admixing said converted thioester with a third oligopeptide, said third oligopeptide comprising a fragment of said desired protein or domain thereof, and having an N-terminal amino acid residue having an unoxidized sulfhydryl side chain and a free amino group that is capable of forming a β -aminothioester linkage with said C-terminal thioester that rearranges to form an amide bond therein between, wherein said admixing is conducted under conditions sufficient to permit the formation of an amide bond between the C-terminus of said second oligopeptide fragment and the N-terminus of said third oligopeptide fragment.

16. The method of claim 15, wherein said N-terminal amino acid residue is a cysteine residue.

17. A synthetically produced protein of greater than about 35 amino acid residues, wherein all of the residues of said protein are linked to adjacent residues via an amide bond.

18. The synthetically produced protein of claim 17, wherein said protein has greater than about 70 amino acid residues.

19. The synthetically produced protein of claim 17, wherein said desired protein is a naturally isolatable protein.

20. The synthetically produced protein of claim 17, wherein said desired protein is a derivative of a naturally isolatable protein that contains one or more cysteine residues that are not found in said naturally isolatable protein.

21. The synthetically produced protein of any of claims 19 or 20, wherein said naturally isolatable protein is a mammalian protein.

22. The method of claim 21, wherein said mammalian protein is a human protein.

23. The method of claim 22, wherein said human protein is a cytokine.

24. A synthetically produced protein of greater than about 35 amino acid residues, wherein all of the residues of said protein are linked to adjacent residues via an amide bond, said protein being produced by the process of ligating together at least two oligopeptide fragments wherein:

(1) said first oligopeptide fragment has a C-terminal thioester; and

(2) said second oligopeptide fragment has an N-terminal amino acid residue having an unoxidized sulfhydryl side chain and a free amino group that is capable of forming a β -aminothioester linkage with said C-terminal thioester that rearranges to form an amide bond therein between; wherein said ligation results in the formation of an amide bond linking said first and second fragments.

25. The synthetically produced protein of claim 24, wherein said N-terminal amino acid residue is a cysteine residue.

26. The synthetically produced protein of claim 24, wherein said protein has greater than about 70 amino acid residues.

27. The synthetically produced protein of claim 24, wherein said desired protein is a naturally isolatable protein.

28. The synthetically produced protein of claim 24, wherein said desired protein is a derivative of a naturally isolatable protein that contains one or more cysteine residues that are not found in said naturally isolatable protein.

29. The synthetically produced protein of any of claims 27 or 28, wherein said naturally isolatable protein is a mammalian protein.

30. The method of claim 29, wherein said mammalian protein is a human protein.

31. The method of claim 30, wherein said human protein is a cytokine.--

REMARKS

Claims 1-7 been cancelled. New claims 8-31 have been added and are now pending. No new matter has been added.

Respectfully submitted,

Nov 8 2000

Date

Donald G. Lewis

Donald G. Lewis, Reg. No. 28,636

THE SCRIPPS RESEARCH INSTITUTE
Office of Patent Counsel
10550 North Torrey Pines Road
Mail Drop TPC-8
La Jolla, California 92037
(858) 784-2937

[X] Attorney or agent of record
[] Filed Under §1.34(a)

- 1 -

SYNTHESIS OF PROTEINS BY NATIVE CHEMICAL LIGATION

SpecificationField of Invention:

The invention relates to methods and intermediates for chemically ligating two oligopeptides end to end with an amide bond. More particularly, the invention relates to methods and intermediates for chemically ligating oligopeptides wherein an unoxidized N-terminal cysteine of a first oligopeptide condenses with a C-terminal thioester of a second oligopeptide to form a β -aminothioester intermediate which spontaneously rearranges intramolecularly to form an amide bond and the ligation product.

Government Rights:

The invention disclosed herein was supported in part by Grants Number R01 GM 48897-01, Number P01 GM 48870-03, and Number GM 50969-01 from the National Institutes of Health. The United States government may have certain rights to this invention.

Background:

Proteins may be synthesized chemically, ribosomally in a cell free system, or ribosomally within a cell. Advances in each of these areas have significantly improved access to many proteins but have also stimulated demand for yet further improvements.

Proteins owe their diverse properties to the precisely folded three dimensional structures of their polypeptide chains. The three dimensional structure of a protein determines its functional attributes. However, at present, it is difficult to predict and/or fully explain the biological properties of a protein from its

- 2 -

three dimensional structure alone. A better understanding of how structure determines the biological properties of a protein can be achieved by systematically varying the covalent structure of the molecule and correlating the effects with the folded structure and biological function. Accordingly, there is an increased demand for enhanced synthetic techniques for synthesizing new proteins and protein analogs.

Techniques derived from recombinant DNA-based molecular biology can be employed to facilitate the expression of proteins in genetically engineered micro-organisms. The use of site-directed mutagenesis, as disclosed by M. Smith (*Angew. Chem. Int. Ed. Engl.* (1994): vol. 33, p 1214), enables the preparation of large numbers of modified proteins in useful amounts for systematic study, e.g., C. Eigenbrot and A. Kossiakoff, *Current Opinion in Biotechnology* (1992): vol. 3, p 333. The use of innovative approaches increases the range of amino acids that can be incorporated in expression systems and promises to significantly extend the utility of biosynthetic modification of the covalent structure of proteins. (C. J. Noren et al., *Science* (1989): vol. 244, p 182 (1989); J. A. Ellman et al., *Science* (1992): vol. 255, p 197.) However, there appear to be limitations inherent to the nature of ribosomal protein synthesis. (V. W. Cornish, et al., *Proc. Natl. Acad. Sci. USA* (1994): vol. 91, p 2910.)

Chemical synthesis of proteins has also contributed to the exploration of the relationship of protein structure to function. Stepwise solid phase synthesis has permitted the *de novo* preparation of small proteins. (T. W. Muir et al., *Curr. Opin. Biotech.* (1993): vol. 4, p 420.) There are also several examples of the use of stepwise solid phase synthesis of whole proteins to explore the molecular basis of biological

- 3 -

function. (M. Miller, et al., *Science* (1989): vol. 246, p 1149; A. Wlodawer, et al., *Science* (1989): vol. 245, p 616; L.H. Huang, et al., *Biochemistry* (1991): vol. 30, p 7402; and K. Rajarathnam, et al., *Science* (1994): vol. 264, p 90.)

Semi-synthesis through the conformationally-assisted religation of peptide fragments can also be employed, in special instances, to study of the structure/function relationship of proteins. (R. E. Offord, "Chemical Approaches to Protein Engineering", in Protein Design and the Development of New therapeutics and Vaccines, J. B. Hook, G. Poste, Eds., (Plenum Press, New York, 1990) pp. 253-282; C. J. A. Wallace, et al., *J. Biol. Chem.* (1992): vol. 267, p 3852. An important extension of the semisynthesis approach is the use of enzymatic ligation of cloned or synthetic peptide segments. (L. Abrahmsen, et al., *Biochemistry* (1991): vol. 30, p 4151; T. K. Chang, et al., *Proc. Natl. Acad. Sci. USA* (1994): in press.) Although the above methodologies have been successfully applied to the synthesis of proteins and protein analogs, T. W. Muir et al., report that there is a continued interest in the wider application of the tools of organic chemistry to the study of proteins (*Curr. Opin. Biotech.* (1993): vol. 4, p 420.)

Stephen Kent et al. recently introduced the chemical ligation of unprotected peptide segments as an improved route to the total synthesis of proteins. (M. Schnlzer, et al., *Science* (1992): vol., 3256, p 221.) Chemical ligation involves the chemoselective reaction of unprotected peptides to give a product with an unnatural backbone structure at the ligation site. Use of unprotected peptides circumvented the difficulties inherent to classical chemical synthesis, viz complex combinations of protecting groups that lead to limited

- 4 -

solubility of many synthetic intermediates, e.g. K. Akaji, et al., *Chem. Pharm. Bull. (Tokyo)* (1985): vol. 33, p 184. In contrast, the chemical ligation technique has allowed us to make good use of the ability to routinely make, purify, and characterize unprotected peptides 50 or more residues in length. Using optimized stepwise solid phase methods the preparation in good yield and high purity of peptides up to 60 residues is routine. In favorable cases, peptides of 80+ residues can be prepared. (M. Schnolzer, et al., *Int. J. Pept. Prot. Res.* (1992): vol. 40, p 180-193.)

The key aspect of the above approach to chemical ligation is the use of a chemoselective reaction to specifically and unambiguously join peptides by formation of an unnatural (i.e. non-peptide) backbone structure at the ligation site. It has permitted the facile preparation of a wide range of backbone-modified proteins, including analogues of protein domains, e.g., ligated 10F3, the integrin-binding module of fibronectin: 95 residues (M. Williams, et al., *J. Am. Chem. Soc.* (1994): in press.) The catalytic contribution of flap-substrate hydrogen bonds in HIV-1 protease has been elucidated by the chemical synthesis of a homodimer of 99 residue subunits of this protein by chemical ligation. (M. Baca, et al., *Proc. Natl. Acad. Sci. U.S.*, (1993): vol. 90, p 11638.) Chemical ligation has also proven to be useful for the routine, reproducible synthesis of large amounts of proteins in high purity with full biological activity (20). (R. C. deLisle Milton, et al., "Synthesis of Proteins by Chemical Ligation of Unprotected Peptide Segments: Mirror-Image Enzyme Molecules, D- & L-HIV Protease Analogs," in Techniques in Protein Chemistry IV, Academic Press, New York, pp. 257-267 (1992).)

Chemical ligation can also be employed for the

- 5 -

straightforward production of protein-like molecules of unusual topology, e.g., four-helix bundle template-assembled synthetic protein (MW 6647 Da) (P. E. Dawson, et al., *J. Am. Chem. Soc.* (1993): vol. 115, p 7263); homogeneous multivalent artificial protein (MW 19,916 Da) (K. Rose, *J. Am. Chem. Soc.* (1994): vol. 116, p 30); artificial neoprotein mimic of the cytoplasmic domains of a multichain integrin receptor (MW 14,194 Da) (T. W. Muir, et al., *Biochemistry*, (1994): vol. 33, pp 7701-7708; and peptide dendrimer (MW 24,205 Da) (C. Rao, et al., *J. Am. Chem. Soc.* (1994): vol. 116, p 6975. The range of proteins accessible by this technique is limited by the size of the synthetic peptide segments.

A useful extension would occur if one had direct synthetic access to native backbone polypeptide chains up to the size of typical protein domains. (A. L. Berman, et al., *Proc. Natl. Acad. Sci. USA* (1994): vol. 91, p 4044.) Chemical ligation would then be employed to string these domains together to explore the world of proteins in a general fashion.

A modular strategy for the total synthesis of proteins has been developed, based on the convergent chemical ligation of unprotected peptides has been disclosed by L.E. Canne, et al. (presented at the Annual Meeting of the Protein Society, San Diego, July 1994). Protein domains (modules) were prepared by chemical ligation of 50-70 residue segments; these domains were then stitched together to give the target protein. Mutually compatible ligation chemistries are required: intra-domain ligation should optimally yield a stable, peptide-like bond; inter-domain ligation will tolerate a wider variation of properties of the structure formed at the ligation site.

Straightforward total chemical synthesis of proteins represents the realization of an important

- 6 -

objective of organic chemistry. It raises the exciting prospect of unrestricted variation of protein covalent structure made possible by general synthetic access, and will give new impetus to exploration of the structural basis of properties such as folding, stability, catalytic activity, binding, and biological action.

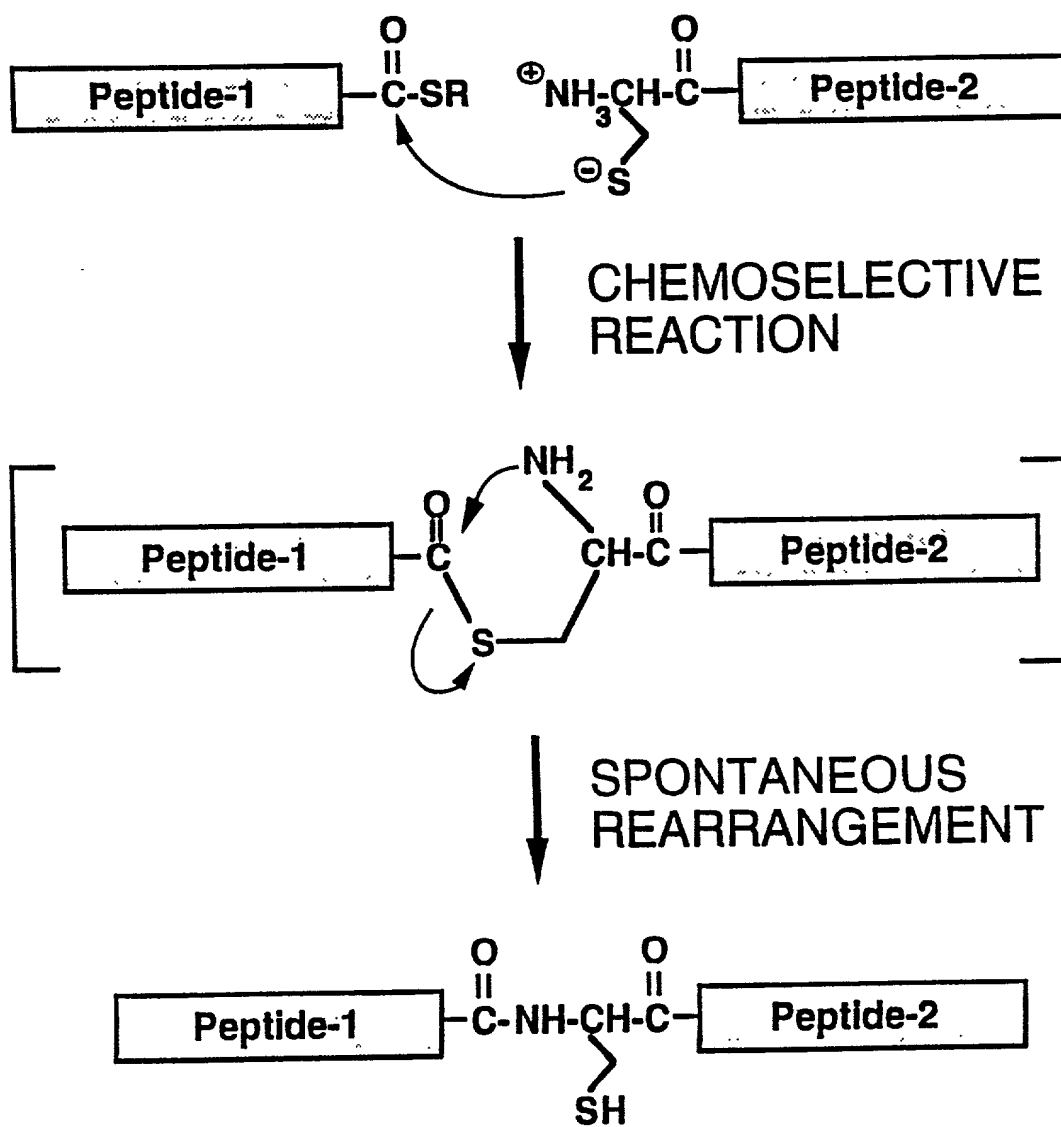
What is needed is a technique of native chemical ligation which combines the formation of a native peptide bond at the ligation site with the advantages of chemoselective reaction of unprotected peptides. This second generation ligation chemistry would significantly increase the size of native backbone polypeptides directly accessible by total chemical synthesis. It could be usefully applied to a wide range of synthetic targets, including proteins of moderate size, and it allows direct access to protein functional domains. Native chemical ligation is a foundation stone of a general modular approach to the total chemical synthesis of proteins. Furthermore, it is compatible with the use of both chemically synthesized peptides and peptide segments derived from other sources.

- 7 -

Summary:

One aspect of the invention is directed to a method of native chemical ligation. The method of native chemical ligation facilitates the chemical synthesis of proteins and large oligopeptide. The principle of 'native chemical ligation' is shown in Scheme 1. The first step is the chemoselective reaction of an unprotected synthetic peptide- α -thioester with another unprotected peptide segment containing an N-terminal Cys residue, to give a thioester-linked intermediate as the initial covalent product. Without change in the reaction conditions, this intermediate undergoes spontaneous, rapid intramolecular reaction to form a native peptide bond at the ligation site. The target full length polypeptide product is obtained in the desired final form without further manipulation. The general synthetic access provided by the method of native chemical ligation greatly expands the scope of variation of the covalent structure of the protein molecule.

One embodiment of the invention provides a method for ligating a first oligopeptide with a second oligopeptide end to end for producing an oligopeptide product. The first and second oligopeptides are admixed in a reaction solution including a catalytic thiol. The catalytic thiol may be an unconjugated mercaptan or a conjugated thiol. Preferred catalytic thiols include benzyl mercaptan, thiophenol, 1-thio-2-nitrophenol, 2-thio-benzoic acid, 2-thio-pyridine, 4-thio-2-pyridinecarboxylic acid, and 4-thio-2-nitro-pyridine. The first oligopeptide includes a C-terminal thioester. The second oligopeptide includes an N-terminal cysteine having an unoxidized sulfhydryl side chain. The unoxidized sulfhydryl side chain of the N-terminal cysteine is then condensed with the C-terminal thioester to produce an intermediate oligopeptide which links the



Scheme 1

- 9 -

first and second oligopeptides with a β -aminothioester bond. The β -aminothioester bond of the intermediate oligopeptide then undergoes an intramolecular rearrangement to produce the oligopeptide product which links the first and second oligopeptides with an amide bond.

Another aspect of the invention is directed to an oligopeptide intermediate which comprises a first oligopeptide segment having a C-terminal thioester, a second oligopeptide segment having a N-terminal cysteine, and a β -aminothioester linkage unit which links the C-terminal thioester and the N-terminal cysteine. The β -aminothioester linkage unit spontaneously rearranges intramolecularly to form an amide bond linking the first and second oligopeptides segments end to end.

Another aspect of the invention is directed to a method for producing an oligopeptide having a C-terminal thioester. The method admixes a resin having a linker with an unoxidized thiol with a Boc-amino acid succinimide ester under reaction conditions to produce a Boc-amino thioester-resin. An oligopeptide is then assembled onto the Boc-amino thioester-resin by stepwise solid phase peptide synthesis. When the oligopeptide is complete, the the Boc-amino thioester-resin is cleaved with HS to produce an oligopeptide having a C-terminal thiol. The C-terminal thiol is then converted to an oligopeptide having a C-terminal thioester.

The oligopeptide thioester (α -COSR moiety) of Scheme 1 can be readily generated from a corresponding oligopeptide thiol ($-\alpha$ COSH) prepared by highly optimized stepwise SPPS on a thioester resin. The thioester resin was prepared by the method of L.E. Canne et al., *Tetrahedron Letters* (1995): vol. 36, pp. 1217-1220, incorporated herein by reference. The method of Canne employs the thioester resin disclosed by Blake and

- 10 -

5 Yamashiro (J. Blake, *Int. J. Pept. Protein Res.* (1981):
vol. 17, p 273; D. Yamashiro, et al., *Int. J. Pept.*
Protein Res. (1988): vol. 31, p 322). Peptide products
were cleaved, purified, and characterized by conventional
methods. (M. Schnolzer, et al., *Int. J. Pept. Protein*
Res., (1992): vol. 40, pp 180-193.)

10 The Yamashiro methodology activates a thiocarboxyl
group on a protected oligopeptide with diaryl disulfides
to give acyl disulfides (Yamashiro et. al. *Int. J.*
Peptide Protein Res. (1992): vol. 31, pp 322-334). These
C-terminal-peptide-acyl-disulfides are highly reactive
electrophilic intermediates which are attacked and
subsequently coupled with an α -amino group on the N-
terminus of a second peptide to form native peptide
15 bonds. The reported coupling yields using 2,2'-dipyridyl
disulfide as the activator of the thiocarboxyl group
afford the desired α -IB-92 product in 45% yield. Overall
yields based on the starting resin for a 3-segment
synthesis of α -IB-92 are reported as 8%, while a 2-
20 segment synthesis gave 11%.

Due to the high reactivity of the diaryl disulfide
bond, the Yamashiro approach requires extensive
protection and deprotection of amino acid residues
present in the peptide molecule. The lysine group for
25 example is protected as a citraconyl derivative because
of the reactive amine functionality. Additionally, an
Msc group or tBOC group is used to protect any terminal
amine functionalities present in the molecule.

30 The invention stated herein does not require the
use of any protecting groups for the coupling of two
oligopeptides because a less reactive (and thus more
chemoselective) thioester electrophile is used instead of
the acyl disulfide moiety (Yamashiro's approach). In the
intermolecular coupling step, this thioester electrophile
35 requires a more nucleophilic sulfhydryl moiety rather

- 11 -

than a free amine. The nucleophilic sulfhydryl moiety can be found on cysteine residues. Since the amino and hydroxyl functionalities are relatively unreactive to the thioester electrophile, a selective coupling of the two unprotected oligopeptides is achieved with the cysteine sulfhydryl moiety. The sulfhydryl group on the cysteine of peptide 2 will first attack the thioester of peptide 1 and form a coupled thioester intermediate. This coupled thioester intermediate is concomitantly attacked by the free α -amino moiety from the cysteine and spontaneously rearranges to form the native peptide bond. Yields are therefore increased by eliminating protection and deprotection steps, since side undesired reactions are reduced (Scheme 1).

The thioester moiety is prepared from a precursor thioacid which is obtained by optimized stepwise solid-phase peptide synthesis on a aminomethyl resin support equipped with a thioester resin linker. The precursor thioacid is subsequently generated in liquid HF at 0 °C in 1 hour (Yamashiro et. al. *Int. J. Pept. Protein Res.* (1988): vol. 31, p 322).

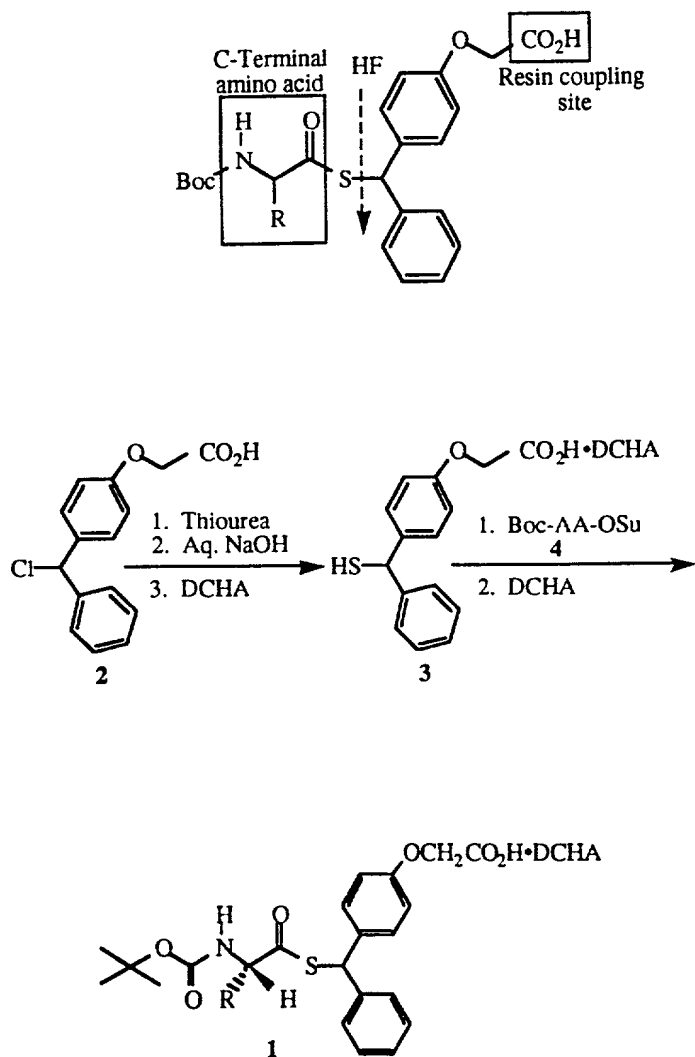
A procedure for the synthesis of the thioester linker with use of a stepwise solid phase peptide synthesis has been reported by Blake (Blake et. al. *Proc. Natl. Acad. Sci. USA* (1981): vol. 78, 4055) and Yamashiro (Yamashiro et. al. *Int. J. Pept. Protein Res.* (1988): vol. 31, p 322). This method is undesirable, however, because it requires the conversion of Boc-amino acid succinimide esters to the corresponding Boc-amino thioacids with hydrogen sulfide. An improved methodology reported herein, utilizes the Boc-amino acid succinimide ester directly and therefore avoids the inconvenience and hazards of hydrogen sulfide gas (Kent et. al. *Tetrahedron Lett.* (1995): vol. 36, p 1217).

In this method (Scheme 2), thiol 3 is generated

- 12 -

from the reaction of chloride 2 (Yamashiro et. al. *Int. J. Pept. Protein Res.* (1988): vol. 31, p 322) with thiourea, followed by hydrolysis of the resulting thiouronium salt in aqueous base. Thiol 3 is a general
5 intermediate which can be reacted with a wide range of commercially available Boc-amino acid succinimide esters to produce the desired thioester linker 1 which is conveniently isolated as the dicyclohexylamine (DCHA) salt.

10 Model studies were undertaken with small peptides to investigate the native chemical ligation approach. To help explore the mechanism of the reaction, the peptide Leu-Tyr-Arg-Ala-Gly- α -COSBzl was reacted with Ac-Cys. The exact mass of the resulting ligation product was
15 determined by electrospray mass spectrometry, and was consistent with a thioester-linked peptide as the ligation product generated by nucleophilic attack of the Ac-Cys side chain on the α -thioester moiety of the peptide. Reaction of Leu-Tyr-Arg-Ala-Gly- α -COSBzl with
20 H-Cys-Arg-Ala-Glu-Tyr-Ser (containing an unblocked α -NH₂ functional group) proceeded rapidly at pH 6.8 (below pH 6 the reaction proceeded very slowly, suggesting the involvement of the ionized thiolate form of the Cys side chain), and gave a single product of the expected mass.
25 This product lacked susceptibility to nucleophiles, and had the ability to form disulfide-linked dimeric peptides, indicating unambiguously the formation of a native amide bond at the ligation site. These studies were consistent with the mechanism shown in Scheme 1, in
30 which the initial thioester ligation product was not observed as a discrete intermediate because of the rapid rearrangement to form a stable peptide bond. Facile intramolecular reaction results from the favorable geometric arrangement of the α -NH₂ moiety with respect to
35 the thioester formed in the initial chemoselective



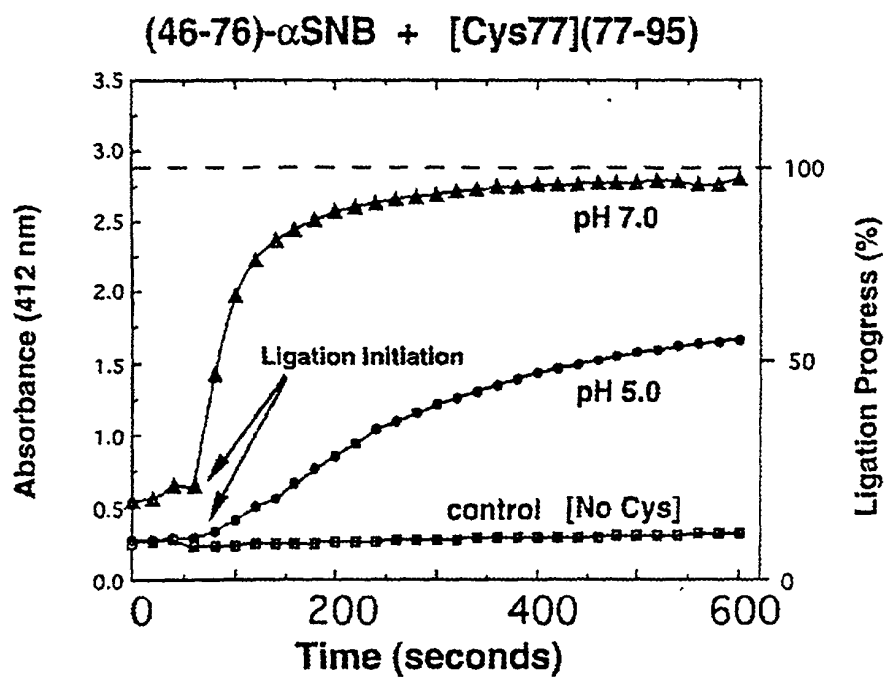
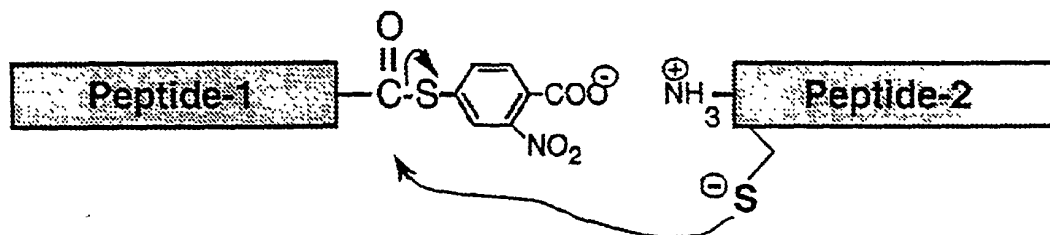
Scheme 2

- 14 -

ligation reaction. Use of such 'entropy activation' for peptide bond formation is based on principles enunciated by Brenner. (M. Brenner, in Peptides. Proceedings of the Eighth European Peptide Symposium H. C. Beyerman, Eds. (North Holland, Amsterdam, 1967) pp. 1-7.) The concept of 'entropy activation' for peptide bond formation has been more recently adopted by D. S. Kemp et al. (*J. Org. Chem.* (1993): vol. 58, p 2216) and by C.-F. Liu, et al. (*J. Am. Chem. Soc.* (1994): vol. 116, p 4149).

Several model peptides have been synthesized by the method of native chemical ligation. The successful synthesis of these model peptides establish that native chemical ligation is generally applicable to peptides containing the full range of functional groups normally found in proteins. Even free internal Cys residues may be present in either of the reacting segments. Internal Cys residues can undergo ester exchange with the peptide- α -thioester component; however, this reaction is unproductive because no rearrangement to the amide bond can occur; the thioester formed is readily reversible and remains a productive part of the reacting system. As disclosed herein, native chemical ligation is limited to reaction at an N-terminal Cys residue. It is important to prevent the side chain thiol of this Cys from oxidizing to form a disulfide linked dimer, because this is unreactive in the ligation. An excess of thiol corresponding to the thioester leaving group was used to keep the Cys residues in reduced form without interfering with the ligation reaction. The amino-terminal peptide segment must be prepared by chemical synthesis to equip it with the necessary α -COSY functionality. Furthermore, for optimal ligation, this component should have an unhindered (i.e. non β -branched) C-terminal amino acid. Solubilizing agents such as urea or guanidine hydrochloride did not interfere with the ligation and

NATIVE CHEMICAL LIGATION - REACTION RATE



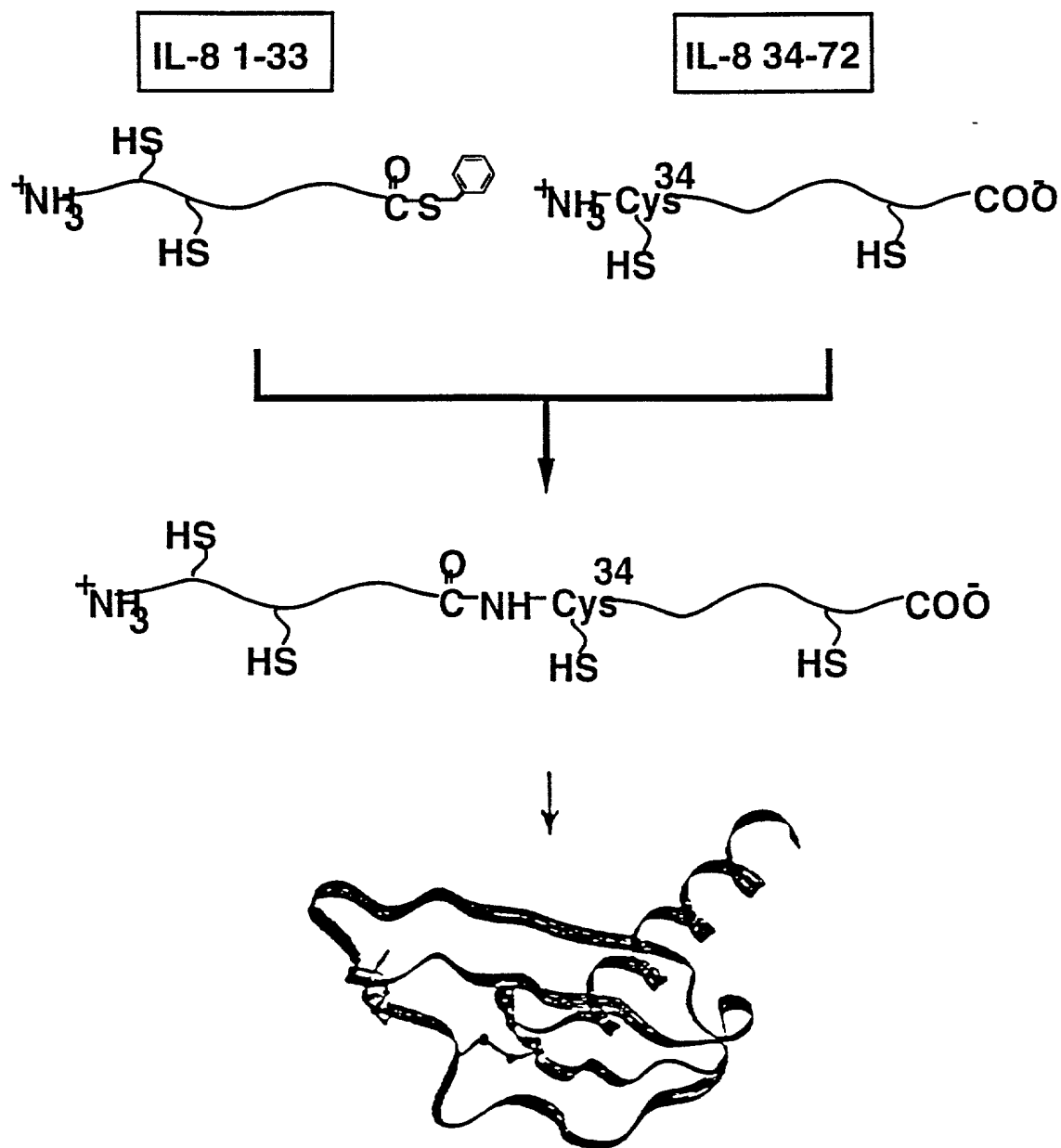
Scheme 5

- 16 -

could be used to enhance the concentration of peptide segments, and thus increase the reaction rate.

Further model reactions demonstrate that the use of better thioester leaving groups results in faster ligation reactions. We applied this observation to the native chemical ligation of peptides from the extracellular domain of a human cytokine receptor (R. D'Andrea, et al., *Blood*, (1994): vol. 83, p 2802.) as shown in Scheme 5. Use of the 5-thio-2-nitrobenzoic acid (-SNB) leaving group, corresponding to the reduced form of Elman's reagent, gave rapid high yield reaction. As described below in connection with Scheme 5, the reaction between the peptide segments was observed to have gone essentially to completion in less than 5 minutes, giving the 50 residue product with a native peptide bond at the site of ligation. Thus, rapid native chemical ligation can be achieved by use of a thioester leaving group with suitably tuned properties.

Application of the native chemical ligation method to the total synthesis of a protein molecule was illustrated by the preparation of human interleukin 8 (IL-8). (M. Baggiolini, et al., *FEBS Lett.* (1989): vol. 307, p 97; I. Clark-Lewis, et al., *J. Biol. Chem.* (1994): vol. 269, p 16075 (1994); I. Clark-Lewis, *Biochemistry* (1991): vol. 30, p 3128; and K. Rajarathnam, et al., (1994): *Biochemistry*, (1994): vol. 29, p 1689.) The 72 amino acid polypeptide chain contains four Cys residues, which form two functionally critical disulfide bridges in the native protein molecule. The total synthesis of IL-8 is shown in Scheme 7. The two unprotected synthetic peptide segments reacted cleanly to give the full length polypeptide chain in reduced form without further chemical manipulation (9). This successful ligation was particularly significant because the 33- and 39-residue IL-8 segments each contained two Cys residues, and



Scheme 7

- 18 -

together encompassed 18 of the 20 genetically encoded amino acids found in proteins. The purified product was folded and oxidized as previously described, to give IL-8 with a mass precisely 4 daltons less than that of the original ligation product, indicating the formation of two disulfide bonds. The properties of this folded product were identical to those of previously studied authentic IL-8 samples. Titration in an assay for neutrophil elastase release demonstrated that the potencies (ED50 = 0.3nM) and maximal responses of the folded, ligated [Ala33]IL-8 and the corresponding molecule obtained by conventional synthesis were indistinguishable and identical to native sequence IL-8. This result unambiguously confirmed the formation of a peptide bond at the ligation site, because the thioester-to-amide rearrangement must have taken place to give the free Cys³⁴ side chain that formed the native disulfide bond (see Scheme 7).

Proteins are usually studied by expression in genetically engineered micro-organisms using the methods of recombinant DNA-based molecular biology. Methods such as site-directed mutagenesis have had a major impact on the ability to prepare large numbers of modified proteins in useful amounts for systematic study. Innovative approaches have increased the range of amino acids that can be incorporated in expression systems and promise to significantly extend the utility of biosynthetic modification of the covalent structure of proteins. However, there appear to be limitations inherent to the nature of ribosomal protein synthesis.

Wieland discloses a method for synthesizing dipeptides using a thioester intermediate. (Wieland et. al. *Liebigs Ann. Chem.* (1953): vol. 580, p 159.) Wieland utilizes the reaction of *S*-glycyl-(or other un-branched aminoacyl-) thiophenols with cysteine. Thus, the

- 19 -

sulfhydryl group on the cysteine residue first attacks the thioester of the S-glycyl-thiophenol and forms a coupled thioester intermediate. This coupled thioester intermediate is concomitantly attacked by the free α -amino moiety from the cysteine and spontaneously rearranges to form the native peptide bond.

A limitation of the Wieland approach is the size of the molecules utilized (only mono-amino acids are coupled to cysteine) and stems from the methodology used for the synthesis of the thioester. To form the thioester, Wieland's approach requires the activation of the terminal carboxylic acid as a mixed anhydride, acid chloride or thioacid. A problem arises if an acidic moiety is present in an amino acid residue such as Asp(D) and Glu(E). In these cases, the Wieland produces an undesired side reaction and therefore requires a complex protecting group strategy, particularly if oligopeptides are synthesized.

The invention described herein eliminates the need for an elaborate protecting group strategy since the oligopeptide-thioester moiety is derived from a precursor thioacid. This precursor thioacid (peptide- α -COSH) is synthesized by a standard stepwise solid-phase peptide synthesis on an aminomethyl resin support, equipped with a thioester resin linker. The precursor thioacid is cleaved from the linker/resin almost quantitatively (99%) in liquid HF at 0 °C for 1 hour.

The thioester peptide (peptide- α -COSR) can be synthesized in two general ways:

(1) Reaction of a crude lyophilized thioacid peptide (peptide- α -COSH) with Ellman's reagent (5,5'-dithiobis-2-nitrobenzoic acid, available from Aldrich company) at pH 5.5 (2.0 equivalents), 6M Guanidine in 100mM Na acetate buffer. This gives the SNB-thioester peptide (peptide- α -COSNB) which is subsequently purified by reversed phase

- 20 -

high performance liquid chromatography (RPHPLC).

(2) Reaction of a crude lyophilized thioacid peptide (peptide- α -COSH) with benzyl bromide at pH 4.0, 6M guanidine and 100mM Na acetate buffer. The benzyl thioester (peptide- α -COSBn) is then purified by RPHPLC.

The conditions stated above, permit the formation of an unprotected oligonucleotide which is equipped with the activated thioester. Subsequent reaction with a second peptide containing a terminal cysteine residue, permits a facile coupling with the formation of a native peptide bond and can generate oligopeptide chains of 100 or more amino acid residues (Scheme 1).

In favorable cases, chemical synthesis has already made important contributions to the exploration of the relationship of protein structure to function. Stepwise solid phase synthesis has permitted the *de novo* preparation of small proteins (14) and there have been several notable examples of the use of this method of total protein synthesis to explore the molecular basis of biological function. Another method that has in special instances allowed chemistry to be applied to the study of proteins is semi-synthesis through the conformationally-assisted religation of peptide fragments. An important extension of the semisynthesis approach is the use of enzymatic ligation of cloned or synthetic peptide segments. Although these methods currently have severe limitations, there continues to be serious interest in the wider application of the tools of organic chemistry to the study of proteins.

Native chemical ligation provides precisely that capability. It combines the formation of a native peptide bond at the ligation site with the advantages of chemoselective reaction of unprotected peptides. This second generation ligation chemistry dramatically increases the size of native backbone polypeptides

- 21 -

directly accessible by total chemical synthesis. It can be usefully applied to a wide range of synthetic targets, including proteins of moderate size, and it allows direct access to protein functional domains. Native chemical
5 ligation is a foundation stone of a general modular approach to the total chemical synthesis of proteins. Furthermore, it is compatible with the use of both chemically synthesized peptides and peptide segments derived from other sources.

10 Straightforward total chemical synthesis of proteins represents the realization of an important objective of organic chemistry. It provides for unrestricted variation of protein covalent structure made possible by general synthetic access, and provides new impetus to
15 exploration of the structural basis of properties such as folding, stability, catalytic activity, binding, and biological action.

20 In an alternative embodiment, the carboxy-terminal peptide segment or protein module can be expressed by standard recDNA means; provided the product contained an N-terminal Cys residue, it could be reacted with the synthetic amino-terminal peptide- α -COSR using the native chemical ligation described here to give a product in
25 which part of the protein had derived from chemical synthesis and part from ribosomal synthesis.

Detailed Description:**Peptide- α -thioacid formation**

5 A typical procedure for the formation and utilization of the thioester resin linker for use in the solid phase synthesis of peptide- α -thioacids is as follows: (Kent et. al. *Tetrahedron Lett.* (1995): vol. 36, p 1217).

10

 4-(α -Mercaptobenzyl)phenoxyacetic acid, dicyclohexylamine (3) Scheme 2. A mixture of 2, formed using the conditions as established by Yamashiro et. al. *Int. J. Pept. Protein Res.* (1988): vol. 31, pp 322-334, (7.5 grams, 27 mmol), thiourea (2.3 g, 30 mmol), and ethanol (100 mL) were heated to reflux (conditions as reported by Koenig et al *J. Org. Chem.* (1958): vol. 23, pp 1525-1530). After 4 hours, conversion to the thiouronium salt was essentially complete as shown by TLC (90:5:5 chloroform:Methanol:Acetic acid). 10N NaOH (30 ml) was added and the reflux continued for 2-3 hours. After cooling to room temperature, the reaction mixture was concentrated *in vacuo* to approximately half the original volume, acidified with concentrated HCl (to pH 2.0), and extracted with ethylacetate (4 x 30 mL). The combined ethylacetate extracts were washed with saturated NaCl (1 x 30 mL) and dried over MgSO₄. The volatile materials were removed *in vacuo*. The resulting oil was dissolved in ethylacetate (100 mL) and any insoluble material filtered. DCHA (dicyclohexylamine - available from Aldrich company), (6.0 mL, 30 mmol) was added to the filtrate with stirring. Within a few minutes, a white solid began to precipitate. Diethylether (150 mL) was added and the suspension cooled at -20 °C for several hours. The resulting white solid was filtered, washed

35

09710633-110800

- 23 -

with Diethylether, and dried under vacuum to give 3 (10.3 g, 23 mmol, 84%): ^1H NMR (CDCl_3): δ 7.30 (m, 7H), 6.82 (d, 2H, $J=8.7$ Hz), 5.39 (br s, 1H), 4.40 (s, 2H), 2.81 (m, 2H), 2.23 (br s, 1H, ex D_2O), 1.88-1.02 (comp m, 20H); FAB MS (cesium ion): calc for $[\text{C}_{27}\text{H}_{37}\text{NO}_3\text{S}, \text{H}^+]$ 456.2572, found 456.2572. Anal. Calcd for $\text{C}_{27}\text{H}_{37}\text{NO}_3\text{S}$: C, 71.17; H, 8.18; N, 3.07; S, 7.04. Found: C, 71.11; H, 8.41; N, 3.08; S, 7.09.

General Synthesis of Boc-amino thioester linker (1), dicyclohexylamine salt (scheme 2). A mixture of 3 (3.67 mmol), Boc-Ala-OSu (available from Novabiochem Corp.) (3.68 mmol), DIEA (diisopropylethylamine-5.74 mmol), dimethylformamide (35 mL) and methylene chloride (4 mL) was stirred at room temperature. After several hours, the initial white suspension completely dissolved to give a clear, colorless solution. After 24 hours, the reaction mixture was poured into 1N HCl (150 mL) and extracted with ethylacetate (4 x 35 mL). The combined ethylacetate extracts were washed with 1N HCl (2 x 30 mL), H_2O (1 x 30 mL), saturated NaCl (1 x 30 mL) and dried over MgSO_4 . Volatiles were removed in vacuo. The resulting oil was purified by flash chromatography (925:50:25 Chloroform:MeOH:acetic acid) to give an oil contaminated with Acetic acid. To remove residual Acetic acid, the oil was dissolved in Chloroform (40 mL) and washed with 0.1 N HCl (7 x 10 mL), saturated NaCl (1 x 10 mL) and dried over MgSO_4 . Volatiles were removed in vacuo to give 1 as an oil. This oil was dissolved in diethylether (10 mL) to which was added dicyclohexylamine (1 equivalents). Hexane (100 mL) was added with stirring to separate the dicyclohexylamine salt of 1 as a thick oil from any unreacted dicyclohexylamine. Solvents were decanted from the oil and the oil dissolved in CH_2Cl_2 (30-40 mL). The resulting solution was concentrated in vacuo

- 24 -

to give the dicyclohexylamine salt 1 as a white foamy solid (scheme 2).

An example of the linkage and synthesis on the resin is as follows: 4-[α -(Boc-Ala-S)benzyl]phenoxyacetic acid
5 (0.80 mmol) is added in 9 ml methylene chloride to 1.00 g aminomethyl-resin (0.40 mmol) and treated at 0 °C with 1.33 mL 0.6 M DCCI (dicyclohexylcarbodiimide) in methylene chloride for 15 min and at 24 °C for 30 minutes. The product is then subjected to standard
10 solid-phase peptide synthesis conditions (Kent et. al. *Tetrahedron Letters* (1995): vol. 36, p 1217; J. Blake, *Int. J. Pept. Protein Res.* (1981): vol. 17, p 273). Once the desired chain is synthesized, the peptide resin (approx. 45 μ mol original load) is treated in 8 mL liquid
15 HF (0.8 mL anisole) at 0 °C for 1 hour. After evaporation with nitrogen, the residue is washed with ethyl acetate. The solid is subsequently stirred in water (approx. 15 mL) at 0 °C while adjusting the pH to 6.0 with solid ammonium bicarbonate. Filtration and
20 lyophilization gives the crude thioacid product which can be further purified by preparative HPLC in 30 mg batches.

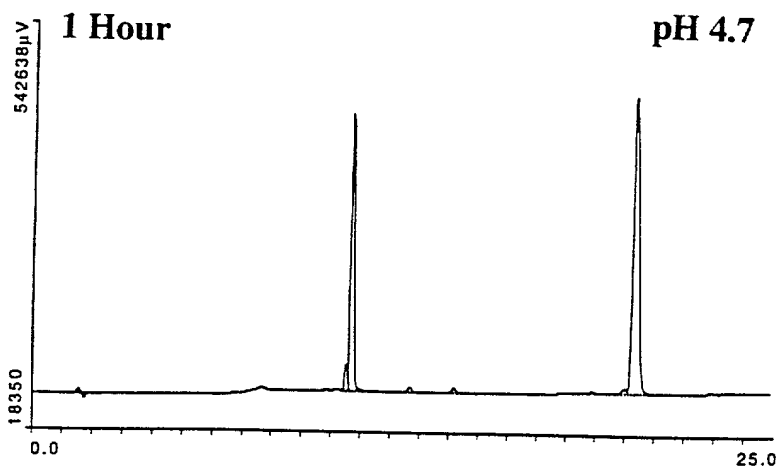
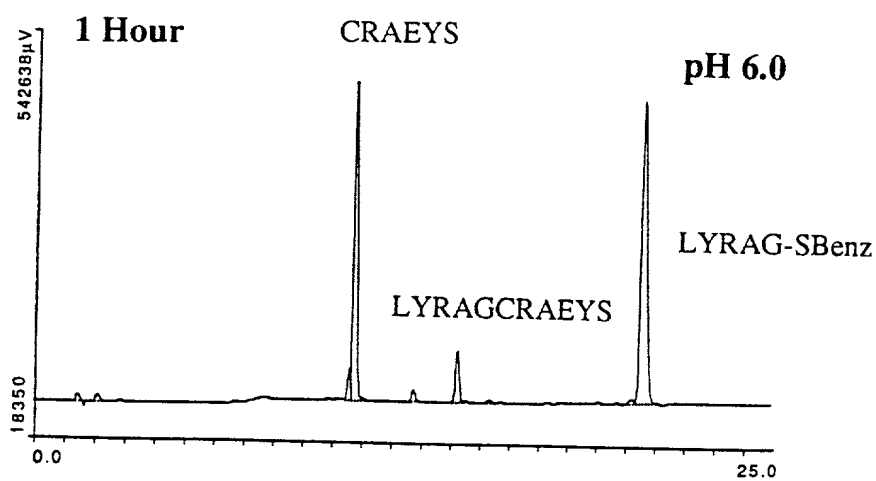
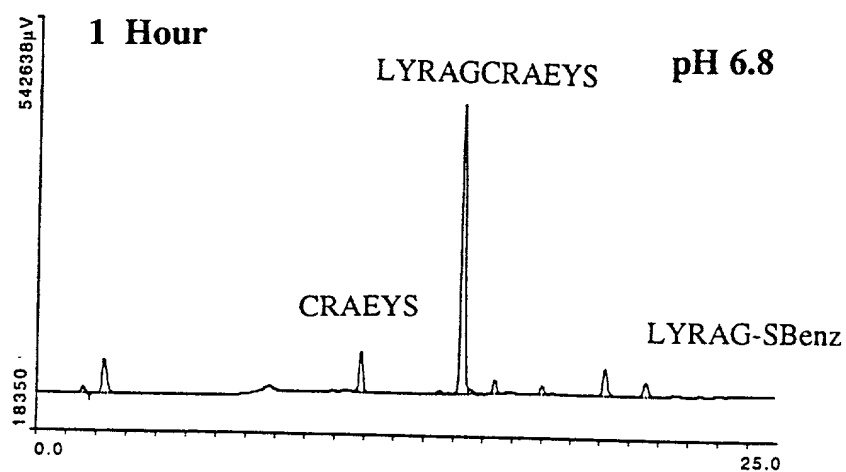
Preparation of the thioester terminal peptide segment

25 The α -COSR thioester peptide can be synthesized in two general ways:

(1) Reaction of a crude lyophilized thioacid peptide with Ellman's reagent (5,5'-dithiobis-2-nitrobenzoic acid, available from Aldrich company) at pH 5.5 (2.0
30 equivalents), 6M Guanidine in 100mM Na acetate buffer. This gives the SNB-thioester peptide which is subsequently purified by reversed phase high performance liquid chromatography (RPHPLC).

(2) Reaction of a crude lyophilized thioacid peptide
35 with benzyl bromide at pH 4.0, 6M guanidine and 100mM Na

-25-



Scheme 3

- 26 -

acetate buffer. The benzyl thioester is then purified by RPHPLC.

5 The conditions stated above, permit the formation of an unprotected oligonucleotide which is equipped with the activated thioester. Subsequent reaction with a second peptide containing a terminal cysteine residue, permits a facile coupling with the formation of a native peptide bond and can generate oligopeptide chains of 100 or more amino acid residues (Scheme 1).

10

09740633-110800

- 27 -

Examples**Example 1**

5 The model peptide Leu-Tyr-Arg-Ala-Gly- α COSH
(Sequence No.: 1) is prepared by optimized stepwise
solid-phase peptide synthesis on an aminomethyl resin.
The thioester resin linker is prepared by a generalized
version as adopted from Kent et. al. *Tetrahedron Letters*,
10 (1995): vol. 36, p 1217; J. Blake, *Int. J. Pept. Protein*
Res. (1981): vol. 17, p 273; D. Yamashiro and C.H. Li,
ibid. (1988): vol. 31, p 322. Once the desired chain is
synthesized, the peptide resin (approx. 45 μ mol original
load) is treated in 8 mL liquid HF (0.8 mL anisole) at 0
°C for 1 hour. After evaporation with nitrogen, the
15 residue is washed with ethyl acetate. The solid is
subsequently stirred in water (approx. 15 mL) at 0 °C
while adjusting the pH to 6.0 with solid ammonium
bicarbonate. Filtration and lyophilization gives the
crude thioacid product which can be further purified by
20 preparative HPLC in 30 mg batches.

 The thioester terminal peptide segment is
subsequently prepared from the thioacid fragment by
chemical synthesis to equip it with the necessary α -COSR
functionality where R is an alkyl group such as benzyl,
25 5-thio-2-nitrobenzoic acid (-SNB), thiophenol, etc. The
use of better thioester leaving groups resulted in faster
ligation reactions. Thus, the model peptide Leu-Tyr-Arg-
Ala-Gly- α COSH (Sequence No.: 1) is first converted to the
thiobenzylester by reaction with benzyl bromide (15
30 equivalents) in 6.0 M guanidine-HCl, pH 4.6, sodium
acetate buffer to form Leu-Tyr-Arg-Ala-Gly- α COSBn
(Sequence No.: 3). The resulting peptide is purified
under standard reversed-phase high-performance liquid
chromatography (HPLC) conditions using approximately 20-
35 45% acetonitrile at 1% per minute; monitored at 214 nm.

- 28 -

For the peptide H-Cys-Arg-Ala-Glu-Tyr-Ser (Sequence No.: 2), solid phase methods allow the preparation of peptides of up to 60 residues in good yield and high purity as described in M. Schnolzer, P. Alewood, D. Alewood, S.B. H. Kent, *Int. J. Pept. Protein Res.* (1992): vol. 40, 180.

First, to explore the mechanism of the reaction, the peptide Leu-Tyr-Arg-Ala-Gly- α COSBn (Bn, benzyl; sequence No.: 3) was reacted with Ac-Cys (containing a blocked α -NH₂ functional group - commercially available from Novabiochem corp.). The exact mass of the resulting ligation product, Leu-Tyr-Arg-Ala-Gly- α COS-CH₂C(NHAc)CO₂H (Sequence No.: 4), was determined by electrospray mass spectrometry and was consistent with a thioester-linked peptide as the ligation product generated by nucleophilic attack of the Ac-Cys side chain on the α -thioester moiety of the peptide.

Finally, the reaction of Leu-Tyr-Arg-Ala-Gly- α COSBn (Sequence No.: 3) with Cys-Arg-Ala-Glu-Tyr-Ser (Sequence No.: 2, containing an unblocked α -NH₂ functional group) proceeded rapidly at pH 6.8 (below pH 6.0, the reaction proceeded very slowly, suggesting the involvement of the ionized thiolate of the Cys side chain at pH 6.8; scheme 3) and gave a single product of the expected mass. The peptides Leu-Tyr-Arg-Ala-Gly- α COSBn + Cys-Arg-Ala-Glu-Tyr-Ser were reacted in 0.1 M phosphate buffer at pH 6.8, 6.0 and 4.7 at 25 °C. After 1 hour, the reactions had proceeded as follows: at pH 6.8 >95%; at pH 6.0 approximately 10%; and at pH 4.7, approximately 1.0% of ligated product Leu-Tyr-Arg-Ala-Gly-Cys-Arg-Ala-Glu-Tyr-Ser (Sequence No.: 5). As observed by HPLC, scheme 3 shows the pH dependence of the reaction after 1 hour and at 25 °C. This product lacked susceptibility to nucleophiles and had the ability to form disulfide-linked dimeric peptides, indicating unambiguously the formation

- 29 -

of a native amide bond at the ligation site.

Another, unpublished model using the 2-thioacetic acid derivative Leu-Tyr-Arg-Ala-Gly-SCH₂COOH (Sequence No.: 6 - ,formed from attack of the thioacid
5 Leu-Tyr-Arg-Ala-Gly-SH (Sequence No.: 1), onto 2-bromoacetic acid in methylene chloride) + Cys-Arg-Ala-Glu-Tyr-Ser (Sequence No.: 2) was ligated at pH 6.8 in 0.2 M phosphate buffer, at 45 °C. After 1.0 hour the reaction had proceeded to 80% as observed in scheme 4 by
10 HPLC. The isolation of oxidation products from the ligated Leu-Tyr-Arg-Ala-Gly-Cys-Arg-Ala-Glu-Tyr-Ser (Sequence No.: 5) and unreacted Cys-Arg-Ala-Glu-Tyr-Ser demonstrated the presence of a free thiol ligation product.

15 The native chemical ligation procedure is generally applicable to peptides containing the full range of functional groups, normally found in proteins. Even free internal Cys residues may be present in either of the reacting segments. Internal Cys residues can undergo
20 ester exchange with the peptide- α -thioester component; however, this reaction is unproductive because no rearrangement to the amide bond can occur, the thioester formed is readily reversible and remains a productive part of the reacting system.

25 The native chemical ligation procedure is limited to reaction at an amino-terminal Cys residue. To prevent the side chain thiol of this Cys from oxidizing to form a disulfide-linked dimer, an excess of thiol
30 corresponding to the thioester leaving group is used to keep the Cys residues in reduced form without interfering with the ligation reaction. In addition, small amounts of low molecular weight thiols such as benzyl mercaptan or thiophenol are added to the coupling reaction mixture to maintain a reducing environment.

09710633-110800

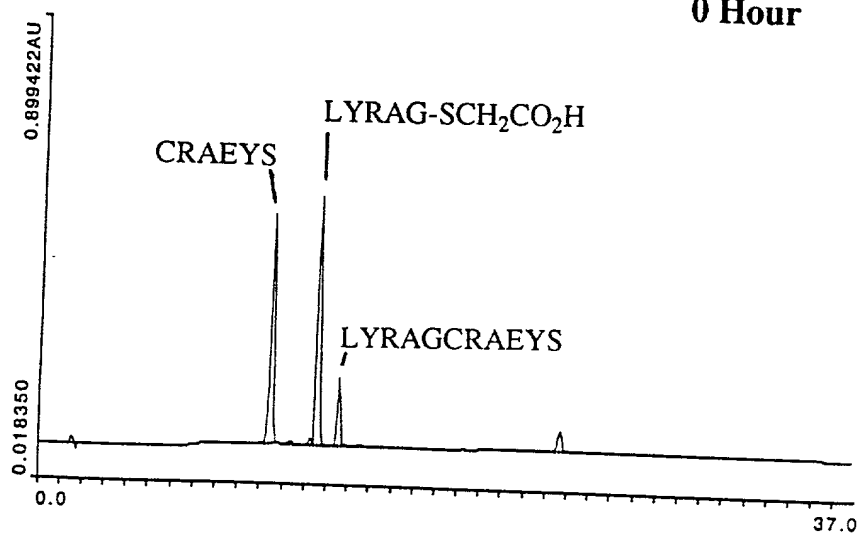
Date: Wed, Feb 2, 1994 3:19 PM
Data: LYRAG-SCH₂COO+CRAEYSI0-001

Sample: pH6.8phos 45degC

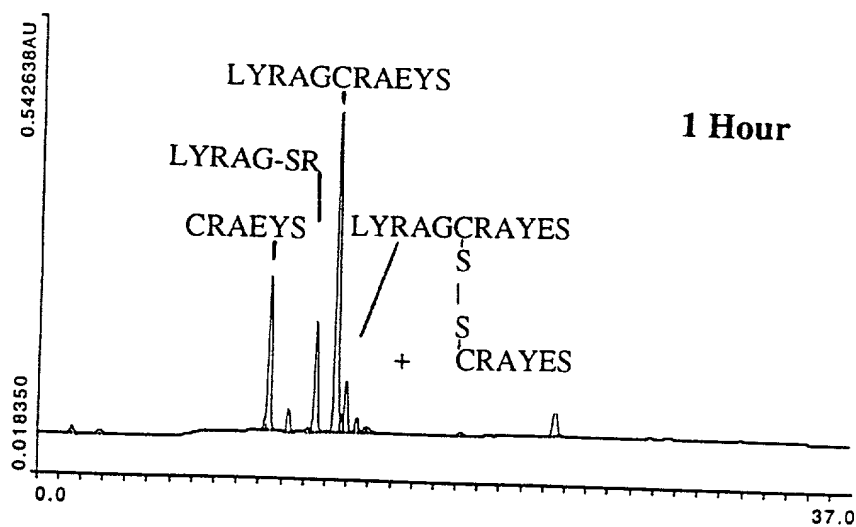
Method: Phil0-67%
Inject Vol: 20
Sampling Int: 0.2 Seconds

Data:

0 Hour



1 Hour



Analysis: Channel A

Peak No.	Time	Type	Height(μV)	Area(μV-sec)	Area%
1	1.506	N1	3161	10542	0.194
2	1.603	N2	9956	54777	1.011
3	2.230	N3	1485	14875	0.274
4	2.833	N	5628	53493	0.987
5	10.276	N1	10621	48495	0.895
6	10.416	N2	177905	958119	17.689
7	11.310	N	27080	145916	2.693
8	12.186	N1	5217	29312	0.541
9	12.376	N2	2379	13576	0.250
10	12.546	N3	127436	729096	13.460
11	13.076	N	2340		
12	13.100	N			

Scheme 4

SUBSTITUTE SHEET (RULE 26)

- 31 -

008077" EESOT/60

The addition of thiols increases the reactivity of the thioester, particularly if the added thiol is a better leaving group than the pre-formed thioester. An example of this observation is when the benzyl ester is converted to a phenyl ester by addition of thiophenol to reaction. Reaction yields and rates are substantially increased. For example, after 7 hours with benzyl mercaptan, the Barnase reaction yielded 25%, while the thiophenol treatment to the same reaction mixture yielded 90%).

Addition of thiols to the ligation mixture also keeps the reaction mixture in a reduced form. This prevents oxidation of the reactive N-terminal Cys residues and when internal Cys residues are present, the thiols reduce the formation of intramolecular disulfide bonds. Additionally, the reducing environment increases the stability of the thioester segment (ligation reactions can proceed overnight with little or no hydrolysis at pH 7.5).

Example 2

A rapid native chemical ligation is illustrated by the synthesis of a peptide segment corresponding to residues 46 to 95 from the external domain of the human IL-3 receptor β -subunit incorporated herein: R.D'Andrea et. al., *Blood* (1994): vol. 83, p 2802.

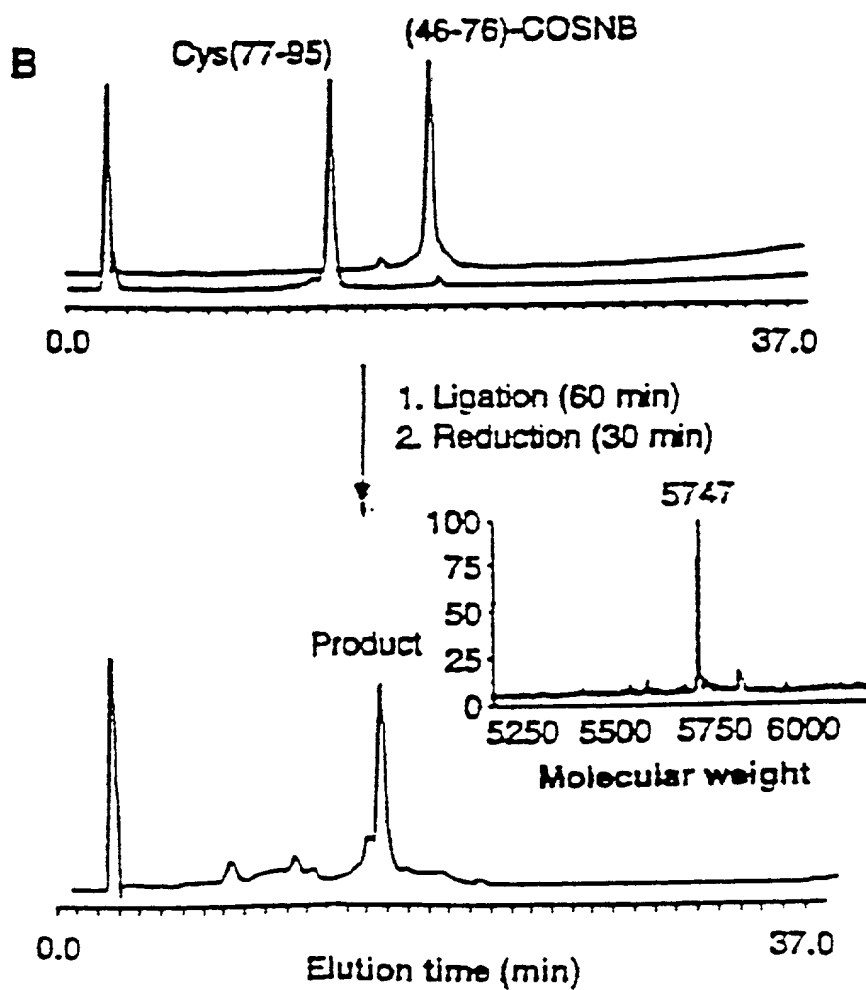
Crude synthetic IL-3 Msc(46-76) α COSH was converted to the 5 thio-2-nitrobenzoic acid ester (-COSNB) by treatment with 5,5'-dithio-bis(2-nitrobenzoic acid) [10 equivalents (eq)] in 8 M urea, pH 4.0, 50 mM ammonium acetate buffer [Msc, 2(methyl-sulfonyl)-ethyloxy-carbonyl (Fluka #69227) protecting group is placed on the N-terminus using 1.1 equivalents (eq.) 2-(methylsulfonyl)ethyl 4-nitrophenyl carbonate, 1.1 eq.

- 32 -

diisopropylethylamine, and .5 M dimethylformamide]. This thioester-containing material was found to be completely stable below pH 6.0, and was readily purified under standard reversed-phase HPLC conditions using approximately 20-45% acetonitrile at 1% per minute and monitored at 214 nm.

As shown in scheme 5, ligation is initiated by adding IL-3 [Cys77](77-95) (prepared by standard solid phase methods Kent et al., *Int. J. Pept. Protein Res.* (1992): vol. 40, 180) to purified IL-3 Msc(46-76) α COSNB at the stated pH and the reaction is monitored by UV [the substituted aryl thiolate leaving group has a characteristic UV absorbtion at 412 nm ($\epsilon_{\text{TNB}}, 412 \text{ nm} = 13,700 \text{ dm}^3\text{mol}^{-1}\text{cm}^{-1}$)]. At pH 7.0, the reaction is essentially complete within 5 min. No reaction is observed when Msc(46-76) α COSNB is exposed to a 10-fold molar excess of Leu-enkephalin (amino-terminal residue, Tyr) at pH 5.0. This control experiment confirms the absolute requirement for an amino-terminal Cys residue at the site of ligation (scheme 5).

Purified IL-3 [Cys77](77-95) (0.98 mM) and IL-3 (46-76) α COSNB (0.9 mM) were reacted in 8 M urea, pH 5.0, 50 mM ammonium acetate buffer at 23 °C (monitored by analytical HPLC 9C18 reversed phase 22.5 to 45% acetonitrile at 0.7% per minute; 214 nm). After 1 hour, the ligation solution is exposed to the reducing agent tris(2-carboxyethyl)phosphine (TCEP) at pH 9.0 and subsequently raised to pH 13.0 to remove the Msc protecting group (Treatment with TCEP was found to aid in



Scheme 6

- 34 -

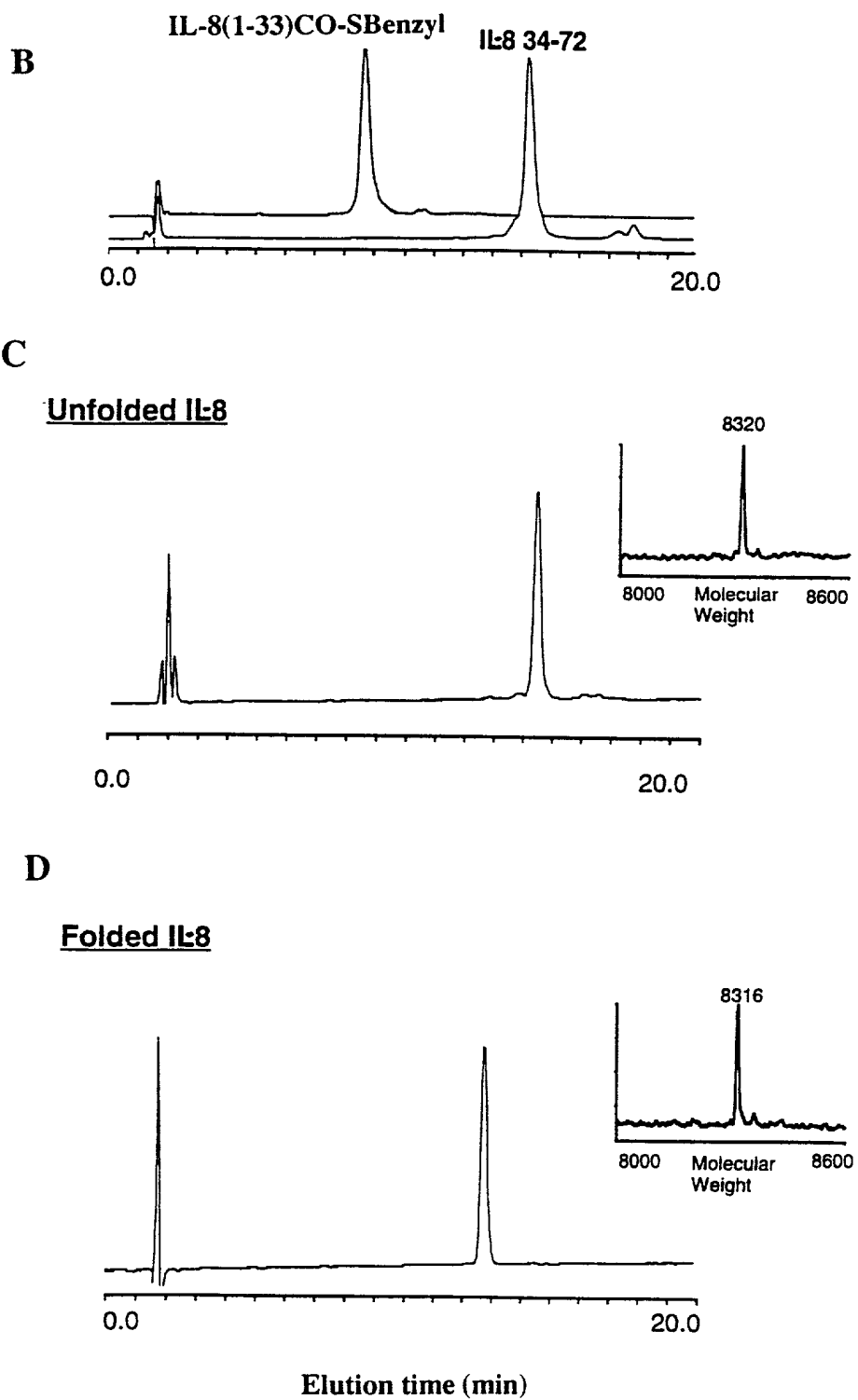
the purification and analysis of product by reducing the thiophenol and benzyl mercaptan disulfide products which had a tendency to co-elute with the peptide products). Scheme 6 shows the progress of the reaction by HPLC; conversion of the starting peptides to the crude product is shown (scheme 6). The 50 residue product has the expected molecular mass by electrospray mass spectrometry [observed, 5747.0 daltons; calculated (average isotope composition), 5747.4 daltons]. The ligation product is shown to be stable at high pH, reducing conditions, and forms an intramolecular disulfide bond. These observations are consistent with the presence of a native peptide bond at the site of ligation.

Analogous methods have required removal of protecting groups (J. Blake et. al., *Int. J. Pept. Protein Res.* (1981): vol. 17, p 273; Kemp et. al. *J. Org. Chem.* (1993): vol. 58, p 2216; Liu et. al. *J. Am. Chem. Soc.* (1994): vol. 116, 4149) or conversion of intermediates to the final form, or both steps. No previous method has allowed the chemical reaction of unprotected peptide segments to directly yield a native backbone final product.

Example 3

The IL-8(34-72) segment (Sequence No. 9) is prepared by optimized stepwise solid-phase methods as described by Kent et al., *Int. J. Pept. Protein Res.* (1992): vol. 40, 180 and yield peptides from 60-80 residues in good yields and high purities. The peptide- α COSH is prepared by optimized stepwise solid-phase peptide synthesis on a aminomethyl resin with a thioester linker. The thioester linker is prepared by a generalized version as adopted from J. Blake, *Int. J. Pept. Protein Res.* (1981): vol. 17, p 273; D. Yamashiro and C.H. Li, *ibid.* vol. 31, 322 (1988). Products are subsequently purified by standard

-35-



Scheme 8

- 36 -

reversed-phase HPLC conditions and characterized by standard methods which include electrospray mass spectrometry.

Crude synthetic segment IL-8(1-33) α COSH (Sequence No.: 7) is converted to the thiobenzyl ester by reaction with benzyl bromide (15 equivalents) in 6 M guanidine-HCl at pH 4.6 in 100 mM sodium acetate buffer. The reaction mixture is purified under standard reversed-phase HPLC conditions and forms the thiobenzyl ester, IL-8(1-33) α COSBn (Sequence No.: 8), (Scheme 7).

The segments IL-8(1-33) α COSBn (Sequence No.: 8), (5.0 mg, 1.3 μ mol) and IL-8(34-72) (Sequence No. 9), (4.8 mg, 1.1 mmol) were reacted in 0.5 ml 6.0 M guanidine-HCl, pH 7.6, phosphate buffer at 23 °C in the presence of benzyl mercaptan (5 ml). After suitable reaction time (48 to 72 hours), a ligation yield of approximately 60% was obtained. The product was purified by standard reversed-phase HPLC as described *via supra* and characterized by electrospray mass spectroscopy.

As shown in scheme 8B, an analytical HPLC spectrum (C_{18} reversed phase; 25 to 45% acetonitrile at 1% per minute; monitored at 214 nm) is shown before the reaction of the synthetic peptide segments IL-8(1-33) α COSBzl and IL-8(34-72).

As shown in scheme 8C, an analytical HPLC spectrum (C_{18} reversed phase; 25 to 45% acetonitrile at 1% per minute; monitored at 214 nm) of the purified ligation product, IL-8(1-72)(SH)₄ (Sequence No.: 10), in fully reduced form. (Inset) Electrospray mass spectrum (raw data displayed as a single charge state): observed molecular mass 8319.8 daltons; calculated molecular mass (average isotope composition), 8319.8 daltons.

As shown in scheme 8D, air oxidation of the purified 1-72 ligation product forms the folded [Ala³³]IL-8 molecule, shown after HPLC purification. The earlier

- 37 -

elution of the folded, disulfide cross-linked native protein compared with the reduced polypeptide is typical (see Lewis et. al *FEB Lett.* (1989): vol. 307, p 97; Lewis et al *J. Biol Chem.* : vol. 269, p 16075; Lewis et. al *Biochemistry* (1991): vol. 30, 3128).

Folding and oxidation conditions: polypeptide at 0.2 mg/ml, 1M guanidine-HCl, pH 8.5 tris buffer, and vigorous stirring in air at ambient temperature (inset). Electrospray mass spectrometry of the oxidized and folded synthetic IL-8 (raw data displayed as a single charge state). Observed molecular mass, 8315.6 daltons; calculated molecular mass (average isotope composition), 8315.8 daltons. (Scheme 8).

Example 4: HIV-1 K41 protease (unpublished conditions)

Ligation reactions are performed in several ways. An optimized procedure for a ligation reaction involving a (5-thio-2-nitrobenzoic acid) SNB thioester is to weigh the two peptides, HIV (1-40)-COSNB (Sequence No.: 11, formed from standard conditions stated herein) and HIV (41-99) (Sequence No.: 12, formed from standard conditions stated herein), as solids in the same reaction vessel and add 6.0 M guanidine HCl pH 6.5 with 100 mM Na acetate (the approximate peptide concentration is 7-13 mg/mL of each peptide).

After 5 min, approx. 2.0 % thiol is added. Two thiol catalysts have been used, viz. benzyl mercaptan (forms the benzyl thioester insitu; Sequence No.: 13) and thiophenol (forms the phenyl thioester insitu; Sequence No.: 14). In the ligation of HIV PR, reaction with benzyl mercaptan gave greater than 60% product yield in 40 hours while the thiophenol gave greater than 80% product yield in 10 hours to form HIV-1 K41 protease (Sequence No.: 15).

Subsequent treatment with TCEP was found to aid in

- 38 -

the purification and analysis of product by reducing the thiophenol and benzyl mercaptan disulfide products which tend to co-elute with peptide products. The product was purified by standard reversed-phase HPLC as described *via supra* and characterized by electrospray mass spectroscopy (Scheme 9).

Example 5. Barnase example (unpublished conditions)

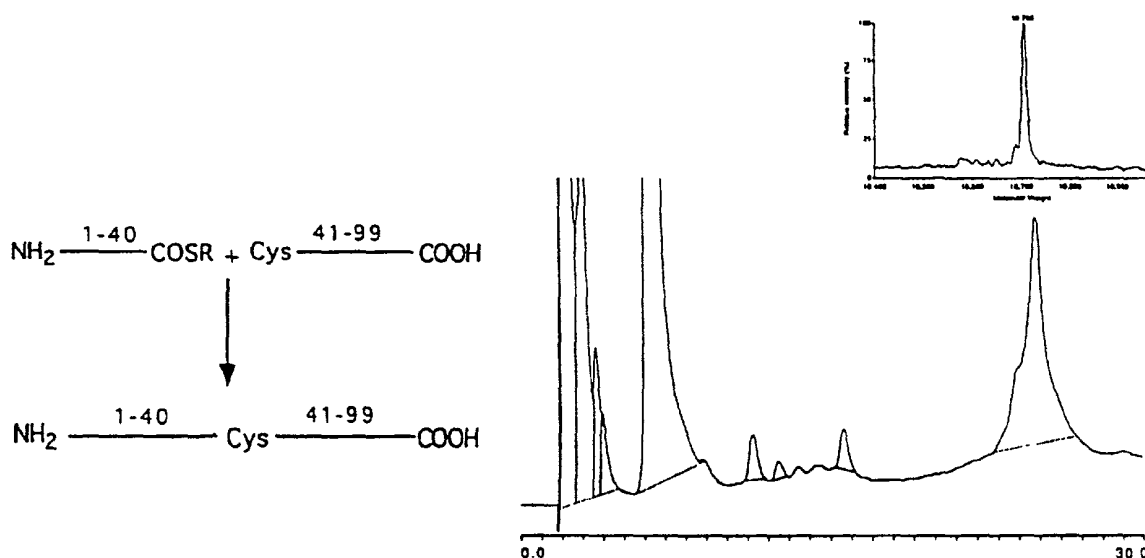
The two peptides Barnase(1-48)-SNB (Sequence No.: 16, formed from standard conditions stated herein) and Barnase(49-110), (Sequence No.: 17, formed from standard conditions stated herein) were weighed as solids in the same reaction vessel and dissolved in pH 7.5 buffer (6M Guanidine 100mM phosphate). Immediately upon dissolving the peptides, 2% benzyl mercaptan (forms the benzyl thioester *insitu*; Sequence No.: 18) or 4% thiophenol (forms the phenyl thioester *insitu*; Sequence No.: 19) was added. After 7 hours, the benzyl mercaptan reaction proceeded 25% and the thiophenol reaction proceeded to > 90% to form Barnase (1-110) (Sequence No.: 20). The product was purified by standard reversed-phase HPLC as described *via supra* (scheme 10).

The addition of thiols increases the reactivity of the thioester, particularly if the added thiol is a better leaving group than the pre-formed thioester. An example of this observation is when the benzyl ester is converted to a phenyl ester by addition of thiophenol to reaction. Reaction yields and rates are substantially increased. For example, after 7 hours with benzyl mercaptan, the Barnase reaction yielded 25%, while the thiophenol treatment to the same reaction mixture yielded 90% to form Barnase (1-110) (Sequence No.: 20).

Addition of thiols to the ligation mixture also keeps the reaction mixture in a reduced form. This prevents oxidation of the reactive N-terminal Cys residues and

Mutant HIV-1 K41 Protease

Synthesized by Native Chemical Ligation



Scheme 9

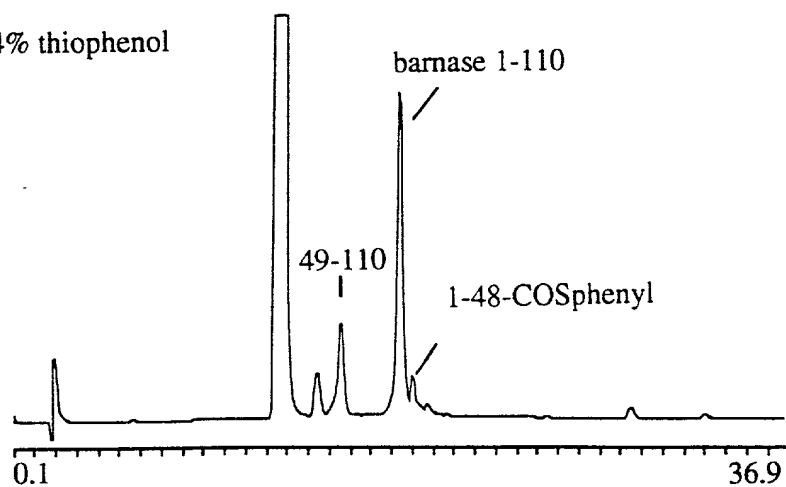
- 40 -

when internal Cys residues are present, the thiols reduce the formation of intramolecular disulfide bonds. Additionally, the reducing environment increases the stability of the thioester segment (ligation reactions
5 can proceed overnight with little or no hydrolysis at pH 7.5).

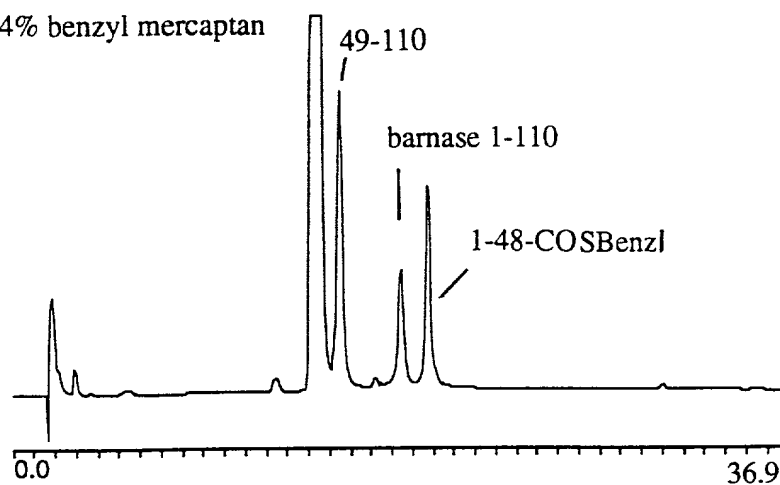
002077 "EE907250

BARNASE K39**1-48COSR + 49(CYS)-110**

7 Hours 4% thiophenol



7 Hours 4% benzyl mercaptan

**Scheme 10**

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: The Scripps Research Institute
- (B) STREET: 10666 North Torrey Pines Road, Suite 220,
Mail Drop TPC8
- (C) CITY: La Jolla
- (D) STATE: CA
- (E) COUNTRY: USA
- (F) POSTAL CODE (ZIP): 92037
- (G) TELEPHONE: 619-554-2937
- (H) TELEFAX: 619-554-6312

(ii) TITLE OF INVENTION: SYNTHESIS OF PROTEINS BY NATIVE CHEMICAL LIGATION

(iii) NUMBER OF SEQUENCES: 20

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(v) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: PCT/US
- (B) FILING DATE: 04-MAY-1995

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- 43 -

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 5
- (D) OTHER INFORMATION: /label = COSH
/note = "Wherein COSH is thioacid."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Leu Tyr Arg Ala Gly
1 5

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Cys Arg Ala Glu Tyr Ser
1 5

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 5
- (D) OTHER INFORMATION: /label = COSBn
/note = "Wherein COSBn is benzyl thioester."

- 44 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Leu Tyr Arg Ala Gly
1 5

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 5
- (D) OTHER INFORMATION: /label = X
/note = "Wherein X is N-acetyl-cysteine-thioester."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Leu Tyr Arg Ala Gly
1 5

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Leu Tyr Arg Ala Gly Cys Arg Ala Glu Tyr Ser
1 5 10

- 45 -

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 5
- (D) OTHER INFORMATION: /label = SCH2COOH
/note = "Wherein SCH2COOH is 2-thioacetic acid."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Leu Tyr Arg Ala Gly
1 5

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 33
- (D) OTHER INFORMATION: /label = COSH
/note = "Wherein COSH is thioacid."

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 1
- (D) OTHER INFORMATION: /label = Msc
/note = "Wherein Msc is
2-methyl-sulfonyl-ethyloxy-carbonyl."

- 46 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Ser Ala Lys Glu Leu Arg Cys Gln Cys Ile Lys Thr Tyr Ser Lys Pro
1 5 10 15

Phe His Pro Lys Phe Ile Lys Glu Leu Arg Val Ile Glu Ser Gly Pro
20 25 30

Ala

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 33
- (D) OTHER INFORMATION: /label = COSBn
/note = "Wherein COSBn is benzyl thioester."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Ser Ala Lys Glu Leu Arg Cys Gln Cys Ile Lys Thr Tyr Ser Lys Pro
1 5 10 15

Phe His Pro Lys Phe Ile Lys Glu Leu Arg Val Ile Glu Ser Gly Pro
20 25 30

Ala

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 amino acids
- (B) TYPE: amino acid

- 47 -

- (C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Cys Ala Asn Thr Glu Ile Ile Val Lys Leu Ser Asp Gly Arg Glu Leu
1 5 10 15

Cys Leu Asp Pro Lys Glu Asn Trp Val Gln Arg Val Val Glu Lys Phe
 20 25 30

Leu Lys Arg Ala Glu Asn Ser
 35

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 72 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
(B) LOCATION: 72
(D) OTHER INFORMATION: /label = SH4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Ser Ala Lys Glu Leu Arg Cys Gln Cys Ile Lys Thr Tyr Ser Lys Pro
1 5 10 15

Phe His Pro Lys Phe Ile Lys Glu Leu Arg Val Ile Glu Ser Gly Pro
 20 25 30

Ala Cys Ala Asn Thr Glu Ile Ile Val Lys Leu Ser Asp Gly Arg Glu
 35 40 45

- 48 -

Leu Cys Leu Asp Pro Lys Glu Asn Trp Val Gln Arg Val Val Glu Lys
50 55 60

Phe Leu Lys Arg Ala Glu Asn Ser
65 70

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 40
- (D) OTHER INFORMATION: /label = COSNB
/note = "Wherein COSNB is 5-thio-2-nitro-benzoic acid ester."

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 27
- (D) OTHER INFORMATION: /label = Xaa
/note = "Wherein Xaa is 2-Aminobutyric acid."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Pro Gln Ile Thr Leu Trp Lys Arg Pro Leu Val Thr Ile Arg Ile Gly
1 5 10 15

Gly Gln Leu Lys Glu Ala Leu Leu Asp Thr Gly Ala Asp Asp Thr Val
20 25 30

Ile Glu Glu Met Asn Leu Pro Gly
35 40

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- [illegible]

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

[illegible]

- [illegible]

[illegible]

- [illegible]

[illegible][illegible][illegible][illegible][illegible][illegible]

- [illegible]

[illegible]

- 50 -

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 40
- (D) OTHER INFORMATION: /label = COSBn
/note = "Wherein COSBn is ??."

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 40
- (D) OTHER INFORMATION: /label = COSBn
/note = "Wherein COSBn is benzyl thio ester."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Pro Gln Ile Thr Leu Trp Lys Arg Pro Leu Val Thr Ile Arg Ile Gly
1 5 10 15

Gly Gln Leu Lys Glu Ala Leu Leu Asp Thr Gly Ala Asp Asp Thr Val
20 25 30

Ile Glu Glu Met Asn Leu Pro Gly
35 40

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 40
- (D) OTHER INFORMATION: /label = COSPh
/note = "Wherein COSPh is phenyl thioester."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Pro Gln Ile Thr Leu Trp Lys Arg Pro Leu Val Thr Ile Arg Ile Gly

- 51 -

1 5 10 15
 Gly Gln Leu Lys Glu Ala Leu Leu Asp Thr Gly Ala Asp Asp Thr Val
 20 25

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 99 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 67
- (D) OTHER INFORMATION: /label = Xaa
 /note = "Wherein Xaa is amino butyric acid."

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 95
- (D) OTHER INFORMATION: /label = Xaa
 /note = "Wherein Xaa is 2-Aminobutyric acid."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Pro Gln Ile Thr Leu Trp Lys Arg Pro Leu Val Thr Ile Arg Ile Gly
 1 5 10 15

Gly Gln Leu Lys Glu Ala Leu Leu Asp Thr Gly Ala Asp Asp Thr Val
 20 25 30

Ile Glu Glu Met Asn Leu Pro Gly Cys Trp Lys Pro Lys Met Ile Gly
 35 40 45

Gly Ile Gly Gly Phe Ile Lys Val Arg Gln Tyr Asp Gln Ile Pro Val
 50 55 60

Glu Ile Xaa Gly His Lys Ala Ile Gly Thr Val Leu Val Gly Pro Thr

009710633-110300

- 52 -

65 70 75 80

Pro Val Asn Ile Ile Gly Arg Asn Leu Leu Thr Gln Ile Gly Xaa Thr

 85 90 95

Leu Asn Phe

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 48
- (D) OTHER INFORMATION: /label = COSNB
/note = "Wherein COSNB is 5-thio-2-nitro benzoic acid ester."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Ala Gln Val Ile Asn Thr Phe Asp Gly Val Ala Asp Tyr Leu Gln Thr

1 5 10 15

Tyr His Lys Leu Pro Asn Asp Tyr Ile Thr Lys Ser Glu Ala Gln Ala

20 25 30

Leu Gly Trp Val Ala Ser Lys Gly Asn Leu Ala Asp Val Ala Pro Gly

35 40 45

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 62 amino acids
- (B) TYPE: amino acid

- 53 -

- (C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Cys Ser Ile Gly Gly Asp Ile Phe Ser Asn Arg Glu Gly Lys Leu Pro
1 5 10 15

Gly Lys Ser Gly Arg Thr Trp Arg Glu Ala Asp Ile Asn Tyr Thr Ser
20 25 30

Gly Phe Arg Asn Ser Asp Arg Ile Leu Tyr Ser Ser Asp Trp Leu Ile
35 40 45

Tyr Lys Thr Thr Asp His Tyr Gln Thr Phe Thr Lys Ile Arg
50 55 60

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
(B) LOCATION: 48
(D) OTHER INFORMATION: /label = COSBn
/note = "Wherein COSBn is benzyl thio ester."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Ala Gln Val Ile Asn Thr Phe Asp Gly Val Ala Asp Tyr Leu Gln Thr
1 5 10 15

Tyr His Lys Leu Pro Asn Asp Tyr Ile Thr Lys Ser Glu Ala Gln Ala
20 25 30

09710633-110600

- 54 -

Leu Gly Trp Val Ala Ser Lys Gly Asn Leu Ala Asp Val Ala Pro Gly
 35 40 45

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 48
- (D) OTHER INFORMATION: /label = COSPh
 /note = "Wherein COSPh is phenyl thio ester."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Ala Gln Val Ile Asn Thr Phe Asp Gly Val Ala Asp Tyr Leu Gln Thr
 1 5 10 15

Tyr His Lys Leu Pro Asn Asp Tyr Ile Thr Lys Ser Glu Ala Gln Ala
 20 25 30

Leu Gly Trp Val Ala Ser Lys Gly Asn Leu Ala Asp Val Ala Pro Gly
 35 40 45

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 110 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- 55 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Ala Gln Val Ile Asn Thr Phe Asp Gly Val Ala Asp Tyr Leu Gln Thr
1 5 10 15

Tyr His Lys Leu Pro Asn Asp Tyr Ile Thr Lys Ser Glu Ala Gln Ala
20 25 30

Leu Gly Trp Val Ala Ser Lys Gly Asn Leu Ala Asp Val Ala Pro Gly
35 40 45

Cys Ser Ile Gly Gly Asp Ile Phe Ser Asn Arg Glu Gly Lys Leu Pro
50 55 60

Gly Lys Ser Gly Arg Thr Trp Arg Glu Ala Asp Ile Asn Tyr Thr Ser
65 70 75 80

Gly Phe Arg Asn Ser Asp Arg Ile Leu Tyr Ser Ser Asp Trp Leu Ile
85 90 95

Tyr Lys Thr Thr Asp His Tyr Gln Thr Phe Thr Lys Ile Arg
100 105 110

09710633-110600

- 56 -

What is claimed is:

1. A method for ligating a first oligopeptide with a
5 second oligopeptide end to end for producing an
oligopeptide product, the method comprising the following
steps:

Step A: admixing the first and second oligopeptides in
a reaction solution including a catalytic thiol, the
10 first oligopeptide including a C-terminal thioester, the
second oligopeptide including an N-terminal cysteine
having an unoxidized sulfhydryl side chain; then

Step B: condensing the unoxidized sulfhydryl side
chain of the N-terminal cysteine with the C-terminal
15 thioester for producing an intermediate oligopeptide
linking the first and second oligopeptides with a β -
aminothioester bond; and then

Step C: rearranging the β -aminothioester bond of the
intermediate oligopeptide of said Step B for producing
20 the oligopeptide product linking the first and second
oligopeptides with an amide bond.

2. A method as described in Claim 1 wherein, in said
step A, the catalytic thiol is selected from the group
25 consisting of unconjugated mercaptans and conjugated
thiols.

3. A method as described in Claim 2 wherein, in said
step A, the catalytic thiol is benzyl mercaptan.

30

- 57 -

4. A method as described in Claim 2 wherein, in said step A, the catalytic thiol is a conjugated thiol selected from the group consisting of thiophenol, 1-thio-
5 2-nitrophenol, 2-thio-benzoic acid, 2-thio-pyridine, 4-thio-2-pyridinecarboxylic acid, and 4-thio-2-nitro-pyridine.

5. A method as described in Claim 4 wherein, in said
10 step A, the conjugated thiol is thiophenol.

6. An oligopeptide intermediate comprising:
a first oligopeptide segment having a C-terminal thioester,
15 a second oligopeptide segment having a N-terminal cysteine, and
a β -aminothioester linkage unit linking the C-terminal thioester and the N-terminal cysteine, said β -aminothioester linkage unit spontaneously rearranging
20 intramolecularly to form an amide bond linking said first and second oligopeptides segments end to end.

09710633-10800
00807-EE907260

- 58 -

7. A method for producing an oligopeptide having a C-terminal thioester, the method comprising the following steps:

5 Step A: providing a resin having a linker with an unoxidized thiol;

Step B: providing a Boc-amino acid succinimide ester; then

10 Step C:: admixing the resin of said Step A and the Boc-amino acid succinimide ester of said Step B under reaction conditions for producing a Boc-amino thioester-resin; then

15 Step D: assembling an oligopeptide onto the Boc-amino thioester-resin by stepwise solid phase peptide synthesis; then

Step E: cleaving the Boc-amino thioester-resin of said Step D with HS for producing an oligopeptide having a C-terminal thiol; and then

20 Step F: converting the oligopeptide having a C-terminal thiol of said Step E to the oligopeptide having a C-terminal thioester.

09710633-110800

PATENT APPLICATION DECLARATION AND POWER OF ATTORNEY

I HEREBY DECLARE THAT:

My residence, post office address, and citizenship are as stated next to my name in PART A of page 2 hereof.

I believe I am the original, first, and sole inventor (if only one name is listed) or an original, first, and joint inventor (if plural names are listed) of the subject matter which is claimed and for which a patent is sought on the invention entitled SYNTHESIS OF PROTEINS BY NATIVE CHEMICAL LIGATION specification of which:

_____ is attached hereto
X was filed on May 4, 1995, as Application Serial No. PCT/US95/05668
and was amended on _____ (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, Sec. 1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, Sec. 119 of any foreign application(s) for patent or inventor's certificate listed in PART B on page 2 hereof and have also identified in PART B on page 2 hereof any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed.

I hereby claim the benefit under Title 35, United States Code, Sec. 120 of any United States application(s) listed in PART C on page 2 hereof and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, Sec. 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, Sec. 1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application.

I hereby declare that all statements made herein of my knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

I hereby appoint the following as my attorneys or agents with full power of substitution to prosecute this application and transact all business in the United States Patent and Trademark Office connected therewith:

Douglas A. Bingham	Reg. No. 32,457	Donald G. Lewis	Reg. No. 28,636
Thomas Fitting	Reg. No. 34,163	Emily Holmes	Reg. No. 40,652

whose mailing address for this application is:

THE SCRIPPS RESEARCH INSTITUTE
10550 North Torrey Pines Road
Mail Drop TPC 8
La Jolla, California 92037

See Page 2 attached, signed, and made a part hereof.

PATENT APPLICATION DECLARATION AND POWER OF ATTORNEY

PART A: Inventor Information And Signature

Full name of SOLE or FIRST inventor Stephen B.H. Kent
Citizenship NZ Post Office Address _____

Residence (if different) _____

Inventor's Signature: _____ Date: _____

Full name of SECOND joint inventor, if any Tom W. Muir
Citizenship GB Post Office Address 500 E 63rd ST (APT# 15B)
NEW YORK, NY 10021

Residence (if different) _____

Second Inventor's Signature: [Signature] Date: 2/10/98

Full name of THIRD joint inventor, if any Philip E. Dawson
Citizenship US Post Office Address 871 Stevens Ave #1314, Solana
Beach, CA 92075

Residence (if different) _____

Third Inventor's Signature: [Signature] Date: 2/11/98

Full name of FOURTH joint inventor, if any _____
Citizenship _____ Post Office Address _____

Residence (if different) _____

Fourth Inventor's Signature: _____ Date: _____

Full name of FIFTH joint inventor, if any _____
Citizenship _____ Post Office Address _____

Residence (if different) _____

Fifth Inventor's Signature: _____ Date: _____

PART B: Prior Foreign Application(s)

Serial No.	Country	Day/Month/Year Filed	Priority Claimed
			<input type="checkbox"/> Yes <input type="checkbox"/> No

PART C: Claim For Benefit of Filing Date of Earlier U.S. Application(s)

Serial No.	Filing Date	Status:
		<input type="checkbox"/> Patented <input type="checkbox"/> Pending <input type="checkbox"/> Abandoned
		<input type="checkbox"/> Patented <input type="checkbox"/> Pending <input type="checkbox"/> Abandoned

See Page 1 to which this is attached and from which this Page 2 continues.

PATENT APPLICATION DECLARATION AND POWER OF ATTORNEY

I HEREBY DECLARE THAT:

My residence, post office address, and citizenship are as stated next to my name in PART A of page 2 hereof.

I believe I am the original, first, and sole inventor (if only one name is listed) or an original, first, and joint inventor (if plural names are listed) of the subject matter which is claimed and for which a patent is sought on the invention entitled SYNTHESIS OF PROTEINS BY NATIVE CHEMICAL LIGATION specification of which:

_____ is attached hereto
X was filed on May 4, 1995, as Application Serial No. PCT/US95/05668
and was amended on _____ (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, Sec. 1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, Sec. 119 of any foreign application(s) for patent or inventor's certificate listed in PART B on page 2 hereof and have also identified in PART B on page 2 hereof any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed.

I hereby claim the benefit under Title 35, United States Code, Sec. 120 of any United States application(s) listed in PART C on page 2 hereof and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, Sec. 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, Sec. 1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application.

I hereby declare that all statements made herein of my knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

I hereby appoint the following as my attorneys or agents with full power of substitution to prosecute this application and transact all business in the United States Patent and Trademark Office connected therewith:

Douglas A. Bingham	Reg. No. 32,457	Donald G. Lewis	Reg. No. 28,636
Thomas Fitting	Reg. No. 34,163	Emily Holmes	Reg. No. 40,652

whose mailing address for this application is:

THE SCRIPPS RESEARCH INSTITUTE
10550 North Torrey Pines Road
Mail Drop TPC 8
La Jolla, California 92037

See Page 2 attached, signed, and made a part hereof.

PATENT APPLICATION DECLARATION AND POWER OF ATTORNEY

PART A: Inventor Information And Signature

Full name of SOLE or FIRST inventor Stephen B.H. Kent
Citizenship NZ Post Office Address 273 Hartford St, San Francisco,
CA 94114
Residence (if different) — as — above —
Inventor's Signature: [Signature] Date: 2-10-98

Full name of SECOND joint inventor, if any Tom W. Muir
Citizenship GB Post Office Address _____
Residence (if different) _____
Second Inventor's Signature: _____ Date: _____

Full name of THIRD joint inventor, if any Philip E. Dawson
Citizenship US Post Office Address _____
Residence (if different) _____
Third Inventor's Signature: _____ Date: _____

Full name of FOURTH joint inventor, if any _____
Citizenship _____ Post Office Address _____
Residence (if different) _____
Fourth Inventor's Signature: _____ Date: _____

Full name of FIFTH joint inventor, if any _____
Citizenship _____ Post Office Address _____
Residence (if different) _____
Fifth Inventor's Signature: _____ Date: _____

PART B: Prior Foreign Application(s)

Serial No.	Country	Day/Month/Year Filed	Priority Claimed
			<input type="checkbox"/> Yes <input type="checkbox"/> No

PART C: Claim For Benefit of Filing Date of Earlier U.S. Application(s)

Serial No.	Filing Date	Status:
		<input type="checkbox"/> Patented <input type="checkbox"/> Pending <input type="checkbox"/> Abandoned
		<input type="checkbox"/> Patented <input type="checkbox"/> Pending <input type="checkbox"/> Abandoned

See Page 1 to which this is attached and from which this Page 2 continues.